Soc. 8

Journal of the AMERICAN STATISTICAL ASSOCIATION

UNIVERSITY OF MICHIGAN

MAR 23 1953

LIBRARY

MARCH 1953

Statistics and the Public Interest ..... *Arynaus Joy Wickens* 1

Data for Measuring the Effectiveness of Public Income-Maintenance Programs ..... *Jacob Fisher* 15

An Appraisal of the 1950 Census Income Data ..... *Herman P. Miller* 23

Delimitation of Economic Areas: Statistical Conceptions in the Study of the Spatial Structure of an Economic System ..... *Rutledge Vining* 44

Labor Productivity in the Soviet Union ..... *Irving H. Siegel* 65

On Probabilities in Bridge ..... *Dan F. Waugh and Frederick V. Waugh* 79

Probabilities of Certain Solitaire Card Games ..... *Robert E. Greenwood* 88

Non-linear Functional Relationship Between Two Variables When One Variable Is Controlled ..... *R. C. Geary* 94

Combination of Neighboring Cells in Contingency Tables ..... *C. C. Craig* 104

Approximating the Mode from Weighted Sample Values ..... *Howard L. Jones* 113

Estimating the Ratio Between the Proportions of Two Classes When One Is a Sub-class of the Other ..... *Jack M. Elkin* 120

The Circular Normal Distribution: Theory and Tables ..... *E. J. Gumbel, J. Arthur Greenwood, and David Durand* 131

BOOK REVIEWS by Jules Backman, John Crawford, Leo A. Goodman, Emil B. Helmes, Millard Hastay, Carl F. Kossack, Margaret G. Reid, and Henry H. Villard 139

PUBLICATIONS RECEIVED 166

RANDOM DIGITS 167

VOLUME 48 NUMBER 261

American Statistical Association

Organized November 27, 1839, Incorporated 1841

The American Statistical Association is a scientific and educational organization. Its membership is not confined to professional statisticians but includes economists, business executives, research directors, government officials, university professors, and others who are seriously interested in the application of statistical methods to practical problems, in the development of more useful methods, and in the improvement of basic statistical data. Engineers, mathematicians, biologists, actuaries, sociologists, psychologists, and representatives of many other professions are included in the membership of the Association.

Annual Dues

Regular membership	$ 8*	Member's subscription to Biometrics	4
Student membership	4	Contributing membership	25
Introductory membership (first dues payment of members under 30 years old)	4	Institutional membership, minimum	100

* Includes $5 for a year's subscription to the *Journal of the American Statistical Association*, and $1.00 for a year's subscription to *The American Statistician*.

Subscription rate, $8 per year. Prices for back issues available on request. A cumulative Index to Volumes 1–34, 1888–1939, may be obtained from the Secretary.

For further information about the Association and membership application forms, write the Secretary, 1108 16th Street, N. W., Washington 6, D.C.

The present Institutional Members are:

American Telephone & Telegraph Company
The Atlantic Refining Company
Armour & Company
Armstrong Cork Company
Chrysler Corporation
The Columbia Gas System, Inc.
Deere & Company
The Dow Chemical Company
Dun & Bradstreet, Inc.
The Equitable Life Assurance Society of the U. S.
Ford Motor Company
General Electric Company
General Foods Corporation
General Motors Corporation
Industrial Surveys Company
International Business Machines
Kennecott Copper Corporation
McCann-Erickson, Inc.
Metropolitan Life Insurance Company
National Analysts, Inc.
National Bank of Detroit
National Industrial Conference Board
Northwest Bancorporation
Philadelphia Electric Company
Public Service Electric & Gas Company
Remington Rand, Inc.
Sears Roebuck & Company
Socony-Vacuum Oil Company, Inc.
Standard Oil Company (Indiana)
Standard Oil Company of New Jersey
Sun Oil Company
United States Steel Corporation
Western Electric Company, Inc.

Published Quarterly by the AMERICAN STATISTICAL ASSOCIATION

Publication Office: 450 Ahnaip Street, Menasha, Wisconsin. Editorial Office: 207 Haskell Hall, University of Chicago, Chicago 37, Illinois. Business Office: 1108 16th St., N.W., Washington 6, D.C. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, section 538, P. L. & R., authorized March 25, 1930. Entered as second class matter at the post office at Menasha, Wisconsin.

Anyone wishing to change his mailing address should allow eight weeks notice. A copy of the address taken from an issue of the periodical should accompany the change of address request.

Bus. Ad.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 48 MARCH 1953 NUMBER 261

ARTICLES

Statistics and the Public Interest . . . . . ARYNESS JOY WICKENS 1
Data for Measuring the Effectiveness of Public Income-Maintenance Programs . . . . . JACOB FISHER 15
An Appraisal of the 1950 Census Income Data . HERMAN P. MILLER 28
Delimitation of Economic Areas: Statistical Conceptions in the Study of the Spatial Structure of an Economic System . . RUTLEDGE VINING 44
Labor Productivity in the Soviet Union . . . . IRVING H. SIEGEL 65
On Probabilities in Bridge . DAN F. WAUGH AND FREDERICK V. WAUGH 79
Probabilities of Certain Solitaire Card Games . ROBERT E. GREENWOOD 88
Non-linear Functional Relationship Between Two Variables When One Variable Is Controlled . . . . . R. C. GEARY 94
Combination of Neighboring Cells in Contingency Tables . C. C. CRAIG 104
Approximating the Mode from Weighted Sample Values . . . . . HOWARD L. JONES 113
Estimating the Ratio Between the Proportions of Two Classes When One Is a Sub-class of the Other . . . . . JACK M. ELKIN 128
The Circular Normal Distribution: Theory and Tables. . . . . E. J. GUMBEL, J. ARTHUR GREENWOOD, AND DAVID DURAND 131

REVIEWS

BOGUE, DONALD J., *State Economic Areas* . . . . RUTLEDGE VINING 44
WALKER, HELEN M., *Mathematics Essential for Elementary Statistics. Revised Edition* . . . . . LEO A. GOODMAN 153
LAMBE, C. G., *Elements of Statistics* . . . . . CARL F. KOSSACK 153
RIOS, SIXTO, *Introducción a los métodos de la estadística. 1ª Parte* . . . . . PAUL R. HALMOS 154
SWEDEN, *Statistiska Centralbyrån, Statistik Tidskrift* 155
NATIONAL BUREAU OF STANDARDS, *A Guide to Tables of the Normal Probability Integral* 155
HOUSE OF REPRESENTATIVES, *Special Subcommittee of the Committee on Education and Labor, Consumers' Price Index* . JULES BACKMAN 156
GERSCHENKRON, ALEXANDER, *A Dollar Index of Soviet Machinery Output, 1927–28 to 1937* . . . . . JOHN CRAWFORD 157
GORDON, ROBERT A., *Business Fluctuations* . . . HENRY H. VILLARD 159
FIRESTONE, O. J., *Private and Public Investment in Canada* . . . . . MILLARD HASTAY 162
KATONA, GEORGE, *Psychological Analysis of Economic Behavior* . . . . . MARGARET G. REID 163
PUBLICATIONS RECEIVED 166
RANDOM DIGITS 167

BUS. ADM.
LIBRARY
HA
1
.A5

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

The Editors welcome the submission of manuscripts for possible publication. They should be typewritten entirely double-spaced, including footnotes, and two copies should be sent to the Editor, W. Allen Wallis, 207 Haskell Hall, University of Chicago, Chicago 37. Books for review should be sent to the same address. Unsolicited book reviews are not accepted, but suggestions of titles for review are welcome.

EDITOR

W. ALLEN WALLIS, *University of Chicago*
ASSISTANT TO THE EDITOR: MARGARET A. LABADIE

ASSOCIATE EDITORS

HOWARD L. JONES
Illinois Bell Telephone Co.
GEORGE M. KUZNETS
University of California
WILLIAM G. MADOW
University of Illinois
PHILIP J. MCCARTHY
Cornell University
I. RICHARD SAVAGE
National Bureau of Standards
C. ASHLEY WRIGHT
Standard Oil Company (N.J.)

ADVISORY PANEL OF FORMER EDITORS

WILLIAM G. COCHRAN (1945–50)
Johns Hopkins University
WILLIAM F. OSBURN (1920–1925)
University of Chicago
FRANK A. ROSS (1926–34, 41–45)
Thetford, Vermont
FREDERICK F. STEPHAN (1935–40)
Princeton University

ERRATA: Readers and authors are urged to submit to the Editor notices of errors found in this or any previous issue. These will be published once a year, in the December issue.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 281 MARCH 1953 *Volume 48*

STATISTICS AND THE PUBLIC INTEREST*

ARYNESS JOY WICKENS
Bureau of Labor Statistics

IN the past few years, statistics in the United States have come to be used as determinants of private and public actions affecting the fortunes of millions of our fellow citizens. As such, they have captured the interest and have become the concern of many who were previously unaware of the existence of statistics.

This development has come so suddenly--and spread into so many fields of activity—that the statistical profession as a whole seems almost unaware of what has been happening. As a profession, we are scarcely prepared, and certainly not organized, to meet the serious responsibilities placed upon us by these new uses of statistics.

It is to these new responsibilities, and the challenge which they present, that I wish to direct your attention this evening. I propose, first, to review the new ways in which statistics are now used; then to discuss the high standards of statistical competence and integrity which these new uses require us to maintain; the new phases of statistical training which they suggest; and the public responsibilities which they impose upon statistical agencies and statisticians.

I need scarcely remind this audience that statistics have come to be one of the great descriptive and analytical tools of modern industrial society, comparable to the other new tools of science. The varied program of this Association at this Annual Meeting is evidence of the fact that statistics are applied today not merely to describe the behavior of man as a social and economic animal, but also to describe plant and animal life, the weather, and, indeed, the shape of the universe.

Statistics of a sort can, of course, be traced back to ancient times, but they have flowered since the industrial revolution. Beginning in

* Presidential Address at the 112th Annual Meeting of the American Statistical Association, Chicago, December 28, 1952.

the 19th century, statistical records were developed to describe the society of that era, and to throw light on its economic and social problems. No doubt they influenced the course of men's thinking then, and even, in some instances, may have led to new policies and new laws; but primarily their uses were descriptive.

Increasingly, in the 20th century, and especially since World War I, statistics have been used to settle problems, and to determine courses of action. In private enterprise, quality control tests now change the production lines of industrial enterprises. New products are developed and tested by statistical means. Scientific experiments turn upon statistics. The management of industry today employs statistics in many ways as a guide to internal operations.

Their use in the sphere of public economic policy-making has expanded almost unbelievably. Beginning with the 1920's, we have seen the development of a whole battery of economic statistics, collected in great detail, by industry, area, and product, and summarized in the familiar indexes and aggregates designed to measure changes in the economic cycle, e.g., the gross national product; the indexes of industrial production; employment and unemployment; financial statistics; price data of many kinds; wages; the volume of trade and commerce—and so on through a long catalogue. As time has gone on, these reports have been compiled in more and more detail, with greater frequency and improved quality.

These statistics come from many sources—from individual business firms and industries and from the professions and their associations; from labor organizations; from schools and colleges; and from private research organizations of many kinds. Because they are often nationwide and necessarily costly, they have been compiled to an increasing extent in the past two decades by the Federal Government. Some of these statistical producing agencies are old, tried, and experienced, and their standards of performance are well known; others are virtually unknown and untested. No license is required to produce statistics. These agencies produce statistics for widely differing purposes and, consequently, some of these summary statistics are finer, more precise instruments than others. But for the most part, these summary economic indicators were intended primarily as analytical tools, for quite broad economic analysis.

For example, they have long been used by the Congress and by executive agencies of the government in making laws and determining policies, as well as by private firms and individuals in making their decisions. As a rule, they have been only one of many factors con-

sidered in reaching a decision, and since they were designed as approximations, they are suitable for just such general purposes as these, where a fairly wide tolerance is acceptable.

World War II, however, made sudden new demands upon statistics for purposes of regulation and control. (There had been some other instances of this kind previously, but they had not been so important.) All at once, detailed economic statistics were in demand for determination of allocations of materials or of shipping space, or for price or wage regulation, or for other control purposes. Those statistics at hand were seized upon, whether or not they were really suitable for control purposes. If no statistics were available, a special-purpose survey was usually hurriedly devised, often with sampling errors of undetermined magnitude. Here, a mistake in classification, in tabulation, or, most important, in the framing of the questions or in the interpretation of the data, had immediate repercussions upon business firms, their employees, and the defense program itself. The statistical profession, it is clear, was quite unprepared for the war emergency. Yet those special wartime statistics were discontinued, with considerable enthusiasm, as soon as controls ended, only to be partially and hastily reassembled for the Defense Program in 1950. Clearly, economists and statisticians have still not provided the public with a design for an adequate, integrated economic intelligence system, based upon sound statistical standards for use in such emergencies. We have not even provided a first blueprint of such a plan.

In still another field, statistics have been used more and more as evidence admissible in court. Both detailed studies relating to particular firms or industries and broad economic indicators appear in attorneys' briefs and exhibits. Statistics are also used by the Federal Trade Commission and the Department of Justice in deciding whether to prosecute a case. The so-called concentration ratios employed in anti-trust prosecutions are an example.

Still more recently, statistics have come to be used as automatic regulators or governors, in the engineer's use of that term. They have been written into law as "trigger figures," determining automatically whether certain actions take place. They have been written into contracts between private firms and individuals, and between governments and private individuals. They have thus come to determine what prices millions of farmers get for their crops, or other millions of workers get for their day's work, or the amount of taxes a business enterprise must pay, or what actual dollar price a producing company receives for a new generator.

Consider for a moment some of these statistical governors. We are familiar with them, but not necessarily with their new uses. The oldest in the United States is the Census, under the Constitutional provision whereby the number of representatives for each State in the House of Representatives shall be determined by the Congress on the basis of a Census of the Population taken once every 10 years. Thus, since the Census of 1950, more representatives come from California and fewer from Pennsylvania. These Census calculations are generally accepted, since this is their primary purpose.

Then, in point of time, there is the parity index, another specially constructed statistical governor. It was devised by Congress after long and bitter agricultural depression, as a yardstick against which to measure farm prices or the farmer's fair share of the national product in comparison with the industry from which he buys and to which he sells. This yardstick—so complex that few economists and too few farmers understand it—is a ratio of indexes of prices paid by farmers to prices received by farmers, hooked back to a pre-World War I base of 1910–14. Today, price supports covering some 40 percent of the sales value of farm commodities are based upon the parity index and parity prices. Congress, having fixed the base period of 1910–14 by law, later amended the Act to add farm wages to the index of prices farmers pay, and to more or less modernize the price relationships as between the individual farm commodities. But for the most part, the parity index continues, with occasional quiet revisions by its custodians in the Bureau of Agricultural Economics, as one of the main factors governing the receipts of millions of farmers.

Then there are several statistical measures of which the Bureau of Labor Statistics is custodian, which have similar uses as "triggers" or governors, only one of which was especially compiled for such purposes.

To take the most recent example first: The Bureau's statistics on housing starts have just been used, under an Act passed by the Congress in the summer of 1952, as a determinant of the time at which the Federal Reserve limitations on housing credit under the so-called "Regulation X" were suspended. These statistics were originally designed as a descriptive measure of the approximate number of new dwellings started in non-farm areas. They were then not calculated on the basis of a seasonally adjusted annual rate, as called for in the Act passed by the Congress, and so a difficult and necessarily rough seasonal calculation had to be made. Moreover, the preliminary form in which they are compiled has sizable potential errors for three months, until final returns come in. Yet the Act provided that they were to be

used as "trigger figures" to suspend these particular credit regulations. It is fortunate that the level of home-building activity in the summer and early autumn of 1952, when this Act became effective, fell so far short of the seasonally adjusted annual rate of housing starts specified in the Act that the decision to suspend Regulation X could be made with statistical confidence. It could easily have been otherwise.

A similar "trigger" use of the Consumers' Price Index was considered when the Defense Production Act of 1950 was being debated in the Congress. In fact, this particular provision once passed the House of Representatives. It provided that price and wage controls during the defense emergency were to go into effect when the Consumers' Price Index had advanced by 5 percent from a given date. However, it was pointed out to members of the Congress that the Consumers' Price Index, being made up of retail prices, would move late and slowly, and that prices of raw materials and industrial products might have sky-rocketed before the Consumers' Price Index had moved by the suggested amount. By then it would have been too late for effective price control measures. Consequently, this particular proposal was dropped.

The venerable Wholesale Price Index of the Bureau of Labor Statistics has also been extensively used as an escalator in business contracts during and since World War II. When this Index was revised early in 1952, it was found that thousands of private contracts involving billions of dollars contained escalator clauses by which the amount to be paid by the purchaser varied automatically with the wholesale index or some component of it. These contracts include government contracts for such equipment as ships, which require several years to construct; many private contracts for heavy equipment or maintenance, such as generators or elevator maintenance; long-term leases on commercial and industrial properties; public utility rates in one State; and a variety of other contracts.

In certain of these contracts in which the materials component varies with wholesale prices of materials, the wage component escalates on average hourly earnings, as compiled by the Bureau of Labor Statistics in connection with its nation-wide monthly survey of manufacturing employment, hours and earnings. This compilation of earnings, while it is very comprehensive, is not intended to be accurate to the last penny.

Another statistical series in the Bureau of Labor Statistics used for escalation purposes is the LIFO (Last In First Out) price index, especially compiled at private expense for the use of department stores in

calculating allowable changes in inventory values for tax purposes, under an agreement with the Bureau of Internal Revenue.

Then there is the well-known use of the Consumers' Price Index in wage contracts, under which wages change immediately and automatically at regular intervals, usually quarterly, whenever there is a specified change in the index. There are also many instances of agreements in which wages or other income payments are subject to renegotiation or reconsideration upon a specified change in the index. In addition to wage contracts, the Consumers' Price Index is used in many other ways as an escalator—for example, in leases for residential buildings, for payments for family maintenance, such as relief payments and alimony payments. These are private contracts negotiated by industry and labor and by private individuals, usually without consultation with the agency which issues the figures. Some of them are fairly long-term, like leases for residential buildings, or in the wage field, like the General Motors five-year contract. Others are for only a year or two.

In addition to these governmental statistics, there are a number of privately compiled yardsticks which serve similar purposes.

Review, now, the effect that these statistics have. A State gains or loses a Congressman; farmers do or do not get loans on their crops; an industrial producer does or does not get a higher price for his products; a tax bill diminishes or increases; wages go up or down. It is no exaggeration to talk of hundreds of millions or even billions of dollars being involved. Let me be specific: Take the case of wage contracts tied to the Consumers' Price Index.

In most of these wage contracts, a change of 1 point or 1.14 points on the present index, which in 1952 has varied around 190 percent of the 1935–39 average, means a change of one cent an hour in wage rates. An increase of one cent an hour for a 40-hour week for 50 weeks out of the year for 3,500,000 workers known, at a minimum, to be under such contracts, means $70,000,000. This does not count millions of other employees whose wages or salaries may follow suit along with the index. Testimony before a Subcommittee of the House Committee on Education and Labor, under the chairmanship of Congressman Steed, which held hearings on the Consumers' Price Index in 1951, indicated that one cent an hour totaled $8,000,000 a year for the General Motors Corporation alone, and $30,000,000 for the railroads, which have over 1,000,000 employees under such contracts.[1] The tables in these wage contracts are set up in brackets of 1 or 1.14 points, by tenths of a point

[1] House Committee on Education and Labor, Subcommittee to Study Consumers' Price Index, 82d Congress, 1st Session, Subcommittee Report No. 2, 1951.

on the index. Thus, at the margins of those wage brackets, if the index rises or falls by one-tenth, hundreds of millions of dollars are immediately involved.

Uses of statistics such as these imply a faith in statistics of which statisticians can be very proud. If statistics and statistical agencies in the United States were not trusted or had not earned a right to public confidence, all this would never have happened. We accept the accolade, but like the Lord Chancellor in *Iolanthe* we say, a little wryly,

> "But though the compliment implied
> Inflates me with legitimate pride,
> It nevertheless can't be denied
> That it has its inconvenient side."

Such responsibility as this is very sobering.

These statistics were usually created, as I have said, for descriptive and analytical purposes, and machinery for their production was geared to more or less leisurely uses by people who were assumed to be reasonably familiar with their construction and their technical limitations. The checking devices, the revision programs, the publicity plans, were not laid out with a view to such awesome uses as these, by a wide, often unknown, and statistically unsophisticated public. The producing agencies have had to change their practices; to speed up their calculations; to guard their results from "leaks"; and greatly to expand their public consulting machinery. As never before, they are operating in a goldfish bowl. They are being advised by numerous committees appointed from the ranks of industry, labor, and technical groups. These advisory committees convey the statistical needs of the groups from which they come to the governmental agencies, and often provide technical advice as well. This type of advisory committee should be continued and expanded if government statistics are to continue to serve broad public uses.

I can say from personal experience that, while these governmental agencies did not seek the exceptional responsibility which the new uses of our statistics have introduced, they have taken the responsibility very seriously. In the opinion of impartial observers, such as the Task Force headed by Professor Frederick C. Mills for the Hoover Commission, the important statistical series produced in government are now administered with competence and honesty and a good deal of skill. That is not to say that they could not be improved, for they could be. But the statisticians who make these statistics know better than anyone that they were never intended to be accurate to the last decimal, as many of their users believe them to be.

By this I do *not* mean that these statistics are not correctly calculated arithmetically. As I have just indicated, I believe they are. Arithmetic is easy to check, and statisticians as a rule do arithmetic well. Most of these major economic indicators are either double-calculated by two different sets of clerks, or they have built-in mechanical checks of various kinds. Of course, there are bound to be occasional errors. No system is infallible. But those errors are accidental. So don't worry about that last decimal! The figures will usually be arithmetically correct, according to the ground rules laid down for the calculation in question. You cannot "rig" a summary figure with 150,000 prices, or millions of people, in it, calculated by a whole battery of individuals, even if someone wanted to—and I do not believe they do.

Rather, I am referring to accuracy in its statistical sense. This depends upon the ground rules. The question is: Was this statistical series *designed*, like the standard meter or the inch, to be accurate to a very small fraction of a point? For most of these general-purpose national economic indicators, the answer, of course, is "No." I repeat—no such accuracy was needed for the purposes for which these statistics were designed. It is not built into the calculations. But, given careful observance of the ground rules, a summary index is usually as correct one time as another, and hence will measure changes from time to time with equal accuracy, except in most unusual circumstances. It must be accepted as it stands, and the public must know that.

The next question, then, is: Did the users choose the proper statistics? Are these indicators really suitable for some of their new uses? Obviously, that question can be answered only with reference to specific cases. But take the major ones: Of the indicators I have named, only the Census, the parity indexes, the LIFO index, and some of the privately-constructed indexes were designed for the specific purposes which they now serve—and none of them is accurate to the last digit, nor is that kind of accuracy required, in my opinion.

However, in many cases, I believe that the other, more general, indicators have also served their new uses well. Most of the statistics which are currently used as automatic "governors" are price indexes, except for the residential construction statistics. They have, in effect, provided insurance to the parties to contracts against wide swings in the price level. Without discussing the more fundamental economic questions involved in these escalator clauses, may I merely say that there has rarely been a time in our recent history when insurance against price changes was so appropriate, or when the price swings were so wide that an approximate indicator was good enough to deal out

rough justice to the parties concerned, provided they chose the appropriate statistical series in the first instance. And for many of the large contracts this has been true.

Now, given these new, more weighty, and more widespread uses of statistics, what do we, as statisticians, owe the trusting public which accepts our products on faith, often with a fateful sentence which begins, "Statistics prove that . . . "

We owe the public, I believe, these things: Arithmetic accuracy; continuing observance of the ground rules for compilation, or a warning if they are to be changed; fairness and impartiality in handling the results; a clear description of the general nature and limitations of statistics, with a simple measure of their accuracy, if it can be measured, and a warning if it cannot. We cannot, and we should not, try to explain every technical detail, or to make statisticians of everyone, any more than all of us should try to understand all the mechanical details of how our automobiles work.

We owe the public also, if we are asked, advice as to whether a certain statistical measure is suitable for a given use or not, and an education in the fact that statistics must and will change. They cannot be expected to be compiled in exactly the same way indefinitely. They get out of repair and go out of fashion, just like other products, if they are not revised from time to time, and so responsible statistical agencies *do* revise them. What would you think of a city consumers' price index which today had the lamp chimneys and high-button shoes of 1900? You may be sure that some of the articles in the indexes of today will look just as peculiar in 25 years. Therefore, anyone making a contract or setting up any other long-term use for a statistical series should be told that he should make provision for periodic shift-overs to revised statistics, if he does not wish to upset the contract. In my opinion, it would be wise also to provide for periodic review of whether circumstances have so changed that other measures—or no statistical measures at all—would be preferable. For example, it should be remembered by those who employ price indexes in contracts that in periods of price stability, which have not been uncommon in our history, such escalators as these indexes provide may either not be needed, or since they are only approximate measures of trends, they may not be accurate to a sufficiently fine point so that their changes have any real significance.

The public should know, too, that there is no real assurance that the statistics they use today will be issued at all 10, 15, or 50 years from now. Although they are authorized by law, there is nothing except custom and the enthusiasm of their sponsors to keep them going,

except, as in the case of the Census, or the parity index, where they have a statutory directive. Other statistics may be discontinued because they have served their primary purpose—although this is unlikely to be true of the major economic indexes—or they may be changed or eliminated because of lack of basic data or of financial or organization support. In general, in government, as in private agencies, budgets are reviewed each year and they are often changed, with considerable danger to the continuity of statistical service. As the Steed Subcommittee of the House Committee on Education and Labor observed in its report on the Consumers' Price Index in 1951:[2]

> "The subcommittee believes that the Consumers' Price Index has become so important that it must be regarded as a fixed charge upon the Government; it should not be subject to yearly fluctuations in budget and at the same time be required to do the same amount of work. . . . Unlike some Government programs, the issuance of a statistic of this type depends almost completely upon continuity of effort. It is impossible to cut the work one year and increase it the next on the same project. . . . "

Therefore, the users of statistics and the profession must be educated to take some responsibility for continuity of statistics where that is important—for today it is not assured.

Finally, statisticians owe the public, I believe, absolute assurance of continued competence, of honesty and fairness in the calculation of these statistics—and indeed of all major statistics, whether they are "trigger figures" or not. Note, I say "assurance." I mean that we, as statisticians, need not merely to *be* competent, fair, and honest, but we need to be able to *prove* to a statistically unsophisticated public that, in fact, our statistics are trustworthy.

This is not merely a public relations problem, although its public relations aspects are very important, but one of fundamental, long-run institutional changes which will assure that statistics continue to serve in the public interest. I believe that the statistical profession and its sister professions in the social and the natural sciences, where statistics are used, should immediately take the initiative and assume responsibility for this public assurance.

Consider what we do *not* have in our profession, which other professions have. There is in our profession, for example, nothing comparable to the Certified Public Accountant, or "Member of the College of Surgeons." There is no such label as a "Certified Public Statistician." There is no certification of the source of statistics like the stamp "U.S.

---

[2] House Committee on Education and Labor, Subcommittee to Study Consumers' Price Index, 82d Congress, 1st Session, Subcommittee Report No. 2, 1951, p. 39.

inspected—choice, good or commercial," and few statistical brand names that mean anything except to the statistically sophisticated. Moreover, for our products, there cannot be standardization, and there is only voluntary labeling of the contents, in footnotes which are often too technical for the public. There is no audit comparable to the audit of a firm of certified public accountants, invited by a corporation to audit its books, in order to assure the stockholders and the public not merely of the accuracy of the arithmetic of its accounting, but also of the validity of the way in which the books are kept.

The time has come, in my opinion, for the statistical profession, in its own interest, to consider devising some means by which these functions can be performed for important statistical series. Such a suggestion implies no doubt of the present validity of existing statistics, nor should it be considered to reflect in any way upon the professional standing or integrity of the statisticians concerned. An accountant does not feel incompetent because he is not a CPA, but for certain work a CPA is required; a company is proud of a "U.S. inspected" label; an accountant does not feel that it reflects on his integrity because an outside accounting firm audits his books. In fact, he welcomes the verification of his procedures and his accuracy. And so should statisticians.

What are those guarantees? What underlies them?

The fundamental guarantee of integrity of any statistics, now or in the future, in public or private agencies, as the Mills Committee[3] observed, lies in the quality and the competence of the people who compile them. While there is no statistical equivalent of admission to the bar, over the past three decades we all know that there has been a great improvement in the quality of statistical work; in basic statistical training in the colleges; and in standards set for statisticians in public and private agencies. There are now advanced degrees in Statistics, with fine training in methodology. In the public service, this Association has helped the Civil Service Commission to define standards for various grades of work. What is now most lacking in statistical training in colleges, I believe, is an adequate appreciation of the fact that statistics are tools to be applied to subject matter. A highly trained mathematical statistician without knowledge of the subject can be quite as dangerous as a subject-matter specialist trying to use statistical methods of which he has little knowledge. Statistical students should therefore be given a well-rounded training in both areas. Advanced statistical train-

[3] The Statistical Agencies of the Federal Government, A Report to the Commission on Organization of the Executive Branch of the Government, by Frederick C. Mills and Clarence D. Long, of the Research Staff of the National Bureau of Economic Research, 1949.

ing should certainly involve practical field work experience wherever that is at all feasible. In view of the fact that statistics today are big business, students should be taught how the important, large-scale statistics are made, and how the individual statistician's role is often to serve as one cog in a big, intricate, mass-production process. Teachers will also do well to give emphasis to an uncodified, but none-theless real, standard of ethics, which demands of the statistician accuracy, impartiality, and integrity in all of his work.

Beyond good training, comes experience in formulating a statistical project, defining the terms, managing a study. The crucial nature of this area of statistical administration is too little recognized. These new uses of statistics make it all important.

With these considerations in mind, I hope that this Association will take positive action to carry out some of the excellent suggestions on training made by Samuel Wilks when he was president of this Association; and at the same time, consider again the possibility of setting up examinations for Certified Public Statisticians, remembering that that label should only be applied to those who have been proved competent in method, in subject matter, and in application.

Finally, I come to the question of reassuring the public that statistics which are so widely used in the public interest are competently and fairly compiled, and are designed to meet public needs adequately. It is time, I believe, for the profession as a whole to share some responsibility for these statistics with those who make them.

Since so many of the important statistical governors and statistics used for policy-making purposes in the United States originate in Federal statistical agencies, we might first consider how this problem could be handled for government statistics. (And here I must emphasize that I speak as a private individual, and not as a representative of a government agency.)

I propose that there be created a new United States Statistical Commission, with responsibility for audit of statistical series, similar to an accounting audit, empowered to put a "certified" label on a statistical product. It should also be charged with investigation of methods, scope, and suitability of statistics, and with making recommendations for future improvements and developmental work. Such a Commission in some respects would be similar to the "Boards of Visitors" which some universities have organized to report to the Board of Trustees; in others like the Inspector General of the Army; and in others like the present Research and Development Board in scientific fields.

The members of such a Commission might be selected from a panel of names suggested by this Association and allied associations dealing with important subjects to be considered by the Commission. Primarily, its membership would be drawn from experts outside government who have had actual experience in such operating problems as are faced in governmental statistics, in addition to knowledge of methodology. Like the United States Central Statistical Board, created in 1933, however, it might also have as members heads of some of the important Federal statistical agencies, serving either ex officio or on a consultative basis, where questions of scope and method were involved, but not, of course, an audit of their statistics.

It should be a continuing body, serving on occasion as required, but with a small full-time staff, and adequate financing, so that our most distinguished statisticians, economists, scientists, and other specialists could reasonably be expected to devote time and attention to its work. Its reports should be made to the highest executive authority, and be generally available to the Congress and to the public.

In making an audit of an important series, such a Commission should rely on sample checks to verify the accuracy of the basic data and the calculations, and should carefully examine the established procedures to see whether they are being followed scrupulously. It should devote a good share of its attention to making suggestions for changes in procedures, if those currently in vogue are not in accordance with the best methods known at the time of review, and to considering whether the data are technically adequate for the uses which are being made of them. After such an examination, the Commission should be in a position to give these statistics a stamp of approval—"Certified Public Statistics"—or to withhold that stamp awaiting improvement —and say why. I, for one, would welcome the constructive suggestions and the heightened public interest that would come from such a review. Private statistical series might also be reviewed by the Commission, on request.

To set up such a Commission will involve both institutional inventiveness and new ways and means. I cannot venture to propose a workable blueprint in all its details, but I can observe that, for government statistics, this is certainly an opportune moment for the creation of such a Statistical Commission. This idea, of course, is by no means original. A Central Commission of Statistics, quite large in size and with somewhat different functions than I am suggesting, has operated with considerable success in the Netherlands since the 1890's, and we might learn from their experience.

The general goals I have outlined will not be easy to accomplish: To lift the level of statistical public relations; to make statistical training more realistic; to devise means for establishing the title of "certified public statistician"; to set up an independent Commission to review and pass upon our most important statistics, governmental and private.

But our profession has made some first steps in these directions. This Association has for many years appointed advisory committees to various governmental agencies, at their request—the Bureau of the Census, the Bureau of Labor Statistics, the Bureau of Mines, and others. These committees have served most usefully in dealing both with program content and with statistical techniques. Similar committees should continue to be appointed in the future. However, they have performed advisory functions of a somewhat different character than are intended for the proposed new Commission. The Committee on Government Statistics and Information Services of 1933, out of which came the U. S. Central Statistical Board and its current successor, the Office of Statistical Standards, was very valuable. Certainly the level of governmental statistics has been raised during the past 20 years, in consequence of its work. The contribution of the report of the Mills Task Force to the Hoover Commission, on government statistics, is such another step. Our Association's Commission on Statistical Standards was created with a much more limited scope, but also in the direction of such functions as are proposed for a Statistical Commission. Its subcommittees which participated with the Social Science Research Council in reviewing the results of the election polls in 1948 and of the statistics underlying the Kinsey Report were both very worth while. This Commission has not been very active, for it is voluntary and part-time, and has no organic relationship to statistical organizations as such.

These few steps, however, are evidence of professional recognition of some of the problems that are before us. Recently, these problems have become much more serious, for statistics have suddenly attained stature beyond our hopes—and perhaps beyond our desires. Their importance goes beyond our present ability to handle the problems which these new uses create with the imagination and authority which the situation requires. As a profession, statisticians must organize to meet this challenge if statistics are to continue to be administered in the public interest.

DATA FOR MEASURING THE EFFECTIVENESS OF PUBLIC INCOME-MAINTENANCE PROGRAMS*

JACOB FISHER
Social Security Administration

The effectiveness of the income-maintenance programs is best appraised in relation to program objectives. These may be considered as falling under two heads—coverage and benefits. Measures for appraising success in achieving coverage and benefit objectives are to be found in the periodic statistical reports and special studies of the administering agencies. Some of the measures acquire significance only when placed in relation to general social and economic data. Since agencies vary in the amount and character of the information they collect and publish, and gaps exist in general social and economic data, the resultant body of knowledge is incomplete, particularly with respect to benefit adequacy.

THE income-maintenance programs with which the present paper deals are public programs administered by public authorities, making money payments to individuals or families based either on need or on the occurrence of an event against which the individual and often his dependents are insured on the basis of past employment or service. More specifically, the programs are public assistance and the social insurance and related programs, i.e., unemployment insurance, disability insurance, old-age insurance, survivors insurance, workmen's compensation, and the compensation and pension programs for veterans. Industrial pensions, privately purchased accident and health policies, annuities and life insurance also help maintain income, but are excluded from consideration here because they present research problems of a somewhat different kind.

The effectiveness of the income-maintenance programs is most appropriately appraised in relation to adequacy of coverage and adequacy of benefit.

The social insurance programs provide protection for specified groups of workers against specified risks. A group is usually identified on the basis of industry, occupation or class of worker, or two of these categories in combination, and sometimes, in addition, by size of establishment, or, as in some workmen's compensation programs, by the

* Based on a paper presented at the session on "The Application of Statistics in Appraising the Effectiveness of the Income-Maintenance Programs," Committee on Statistics in the Social Sciences, American Statistical Association, Boston, December 28, 1951. Opinions expressed in this paper are those of the author and do not necessarily reflect the official views of the Social Security Administration.

hazardous nature of the employment. Further, protection under the social insurance programs is restricted to persons regularly attached to the employments covered, reflecting a general policy to exclude persons engaged only briefly or intermittently in jobs under the system.

The coverage of the public assistance programs is defined in the eligibility conditions. Need, variously interpreted, is common to all of them, but they differ otherwise from program to program with respect to the age classes covered, the nature and severity of the disabilities recognized as bringing the applicant within the scope of the program, and the parental status of the children who could qualify for aid.

The purpose of the benefit in social insurance and related programs is generally, but not always, the partial replacement of a wage loss. It is usually designed to cover the "basic" requirements of the beneficiary and, with the exception of veterans' benefits, is related in amount to previous earnings. The relative emphasis given these goals varies. The proportion of wage loss compensated tends to be higher in the short-run programs, such as unemployment insurance and temporary disability, than in the retirement, survivor, and other long-run programs; and it is higher in the programs with larger contribution rates, and in programs with a more specialized coverage.

The purpose of the public assistance payment is to meet individual need. Wide divergence exists in the interpretation of need. The basic procedure in the determination of the assistance amount, not always honored in practice, is the comparison of an individual's requirements and resources, the assistance payment representing the amount needed to bring total resources up to the requirements recognized in the assistance standard. States vary, however, in their assistance standards; and even within a State, differences may occur in the determination of need. While the States have made progress toward establishing a defined content of requirements to which local administering units must conform in the determination of the assistance amount, the identification and appraisal of requirements and resources are, to a considerable extent, still largely discretionary with the local agency.

What measures are available for gauging the extent to which the income-maintenance programs achieve their goals?

One may turn first to the statistical series maintained by the agencies administering the programs. Special studies constitute another resource, as does that branch of inquiry concerned with the analysis of the legal provisions of the program, the rules and regulations issued by the administering agency, and the decisions of appeals bodies. Not to

be overlooked, either, are the uses to which data on the general characteristics of the population, or of selected groups in it, can be put, since often what is known about the beneficiaries of a program acquires significance only when placed in relation to the demographic universe to which these individuals belong. Although the inventory which follows consists largely of program statistical series, these are not the only source of information for measures of program effectiveness.

Since all earners and their dependents are presumably in need of income security in old age, and the dependents are all presumably in need of survivorship protection, the extent to which the Federal old-age and survivors insurance program and the special systems for government and railroad workers cover all employments is one measure of the coverage of the groups at risk. This may be calculated by reference either to aggregate earnings or to persons employed.

Data on earnings and average number engaged in employments under the Federal old-age and survivors insurance program, the railroad retirement system, and the Federal, State, and local government retirement systems are available on an annual and for some programs on a quarterly and monthly basis.[1] The relative protection of these programs may be measured by the proportion which earnings in covered employments represent of total earnings in the Department of Commerce national income series; or by the proportion which the number in covered employments constitutes of the total number reported in the Census Bureau's monthly report on the labor force. (Note that not all State and local government workers are covered by a retirement system.)

In the Federal old-age and survivors insurance and railroad programs, data have been developed as well on all persons in the covered employment during some part of the year. Because of the constant movement between covered and noncovered employment, more persons are likely to have some earnings credits under a given program during the year than in any one month. During 1950, for instance, 48 million persons earned wages in employments covered under the old-age and survivors insurance program, but average monthly covered employment was 35 million. For some purposes such during-the-year employment figures are better than the during-the-week figures, and may be used in relation to total during-the-year employment as estimated by the Census Bureau and the Social Security Administration. Unlike the case in the average monthly employment estimates,

[1] Sources for statistical series referred to are listed in the Note below.

during-the-year data cannot be added together because some individuals appear in several series.

Eligibility for retirement or survivor benefits in the old-age and survivors insurance and railroad systems is conditioned upon "insured status" which measures attachment to covered employment, or years of service. A more refined measure of the protection afforded than the number of persons in covered employment is the number with insured status or with service credits sufficient to make the worker or his survivors eligible for benefit in case the event insured against should occur. Annual data on the number of earners with insured status are available for the old-age and survivors insurance and railroad programs, as well as data on the number of years with earnings credits or service. These data may be related, for analytical purposes, to data on total employment, after adjustment for the number of insured persons who have left the labor force but still retain insured status.

The employment and insured status data and the data on years with credited wages or services may be had by age, sex, industry (occupation in the railroad retirement program), and for some items, by color, making possible analysis of the bearing of these factors on the extent of protection. The 1 per cent continuous work history sample of the Bureau of Old-Age and Survivors Insurance has yielded several studies of the effect of worker mobility and changes in economic conditions on patterns of covered employment for different age, sex, and industry groups.

Measures of the coverage of extended disability programs may be found in the payroll and employment series of the railroad retirement system and of Federal, State, and local governments, to which reference has already been made. Some earnings and employment series are also available to measure the coverage of the temporary disability programs in the 4 States with such legislation and in the railroad industry. Analysis of the extent to which the disability insurance programs cover all persons presumably in need of such protection may be made by reference to national earnings and employment data issued by National Income Division of the Commerce Department and by the Census Bureau.

The extent of protection against the risk of unemployment is reflected in comparisons which may be made between earners and earnings in covered employment and total wage and salary workers and their earnings. Data on total payrolls and payrolls subject to unemployment insurance contributions, by State, are available on an annual and quarterly basis, while data on the number of workers in covered

employment during an average pay period are published on a monthly and during-the-year basis. An annual series also exists on the number eligible for benefits on the basis of covered earnings during the year, i.e., the number with insured status. Similar data are issued for the railroad unemployment insurance program.

Information on the coverage of workmen's compensation programs is distinctly limited. Until the past decade the emphasis in statistics in this field was upon accidents and relatively little attention was paid to the insurance protection provided. The Social Security Administration annually prepares estimates of covered payrolls for the country as a whole, but estimates are available for a few individual States only and these are in some instances incomplete because limited to particular types of insurers (i.e., private carriers, self-insurers, State fund). Variations among the States in the employments covered, and in the kinds of insurers permitted, and the elective character of the law in some States, make for considerable differences in the information collected and in the comparability, from State to State, of what are presumably the same items.

Another group of coverage questions relates to the extent to which the income-maintenance programs are providing income to persons who have experienced an income loss, i.e., the nonworking aged, the widows and young children of deceased workers, the nonworking disabled, and the unemployed.

Estimates of varying degrees of accuracy and currency are available for the size of each of these groups. The number of aged nonearners in the country may be obtained from the Census Bureau's monthly report on the labor force, while the number and labor force status of widows and their dependent children may be estimated from that Bureau's annual estimates of the family composition and marital status of the population. The number of unemployed is readily obtainable from the same monthly report on the labor force. Estimates of the number of nonearner disabled persons in the working years of life, by age and sex and duration of incapacity, have been developed on the basis of special studies made in conjunction with the Census Bureau's monthly population survey.

The extent to which the social insurance and related programs provide protection against a given income loss may be measured by relating the number of beneficiaries in each of these classes to the total number experiencing the loss, after adjustments are made for differences in definition. The appropriate program data are available to a greater or lesser extent for all programs with the exception of work-

men's compensation (available in some States, however). For some programs, the data come with age, sex, race, industry or occupation detail. In some of the disability insurance programs there is detail also on the kind and duration of disability; and in the temporary disability program, the number of spells of illness in the benefit year. These details are of importance in the analysis of the coverage effectiveness of the program for particular groups of workers. The relative restrictiveness of the disqualification provisions of State unemployment insurance laws is reflected in the ratio of disqualifications to the estimated number of spells of unemployment. Data on the dependents of beneficiaries in relation to the provisions for dependents in the benefit schedule contribute to the appraisal of the completeness of protection with respect to this aspect of the program.

The effective coverage of the public assistance program can be measured only by reference to the total number of needy persons in the community. The development of measures of need has, however, long defied our ingenuity and for good reason. Need is relative and its recognition varies from place to place and over time with living standards, fiscal resources and other factors. The Department of Commerce issues annual estimates of the per capita income of the individual States; these are at best only rough measures of State differences in need. Distributions of family income by county will be available shortly from the 1950 census; when analyzed by family size they should provide valuable indicators of the relative number of families, county by county and State by State, living below a given income standard. Their usefulness in measuring need will be limited however by the time lag between the collection and publication of the data, differences from place to place in living costs, and by the different meanings which will be attached in each county to the significance of the data.

Even if we could agree on a uniform standard for determining need, it would take nothing short of a house-to-house canvass to ascertain the number of persons in the community with incomes currently below the given standard. Public assistance must be applied for: the number of needy persons who might qualify but do not apply for public assistance is not known. What can be measured is the proportion of public assistance recipients in the total population, or in such classes as the aged. In June 1952, for instance, old-age assistance recipients per 1,000 persons 65 years and over varied from 45 in the District of Columbia to 631 in Louisiana. Recipient rates indicate the extent to which population groups are dependent in whole or part upon public assistance. Whether they also measure the extent to which need is

being met through public assistance will depend, among other things, on the scope of the public assistance law, its interpretation by the public assistance agency, and public sentiment.

Monthly data are available on the number of recipients of each type of assistance. Recipient rates are published semi-annually; the population base in old-age assistance is the number of persons 65 years of age and over; in aid to dependent children, the number of children under 18 years; and in aid to the blind, in aid to the permanently and totally disabled and general assistance, the total population. The comparability of State rates in the aid to the blind and disability assistance programs is affected by differences of unknown magnitude in the relative number of blind and disabled persons, State by State. Special studies have developed, in addition, information on the personal and family characteristics of recipients, the labor force history of old-age assistance recipients, the types and duration of disability encountered in the aid to the blind and aid to the permanently and totally disabled programs and the parental status and reasons for dependency of children receiving aid to dependent children. The periodic data and the special studies throw valuable light on the influence of differences in State eligibility standards and in the incidence of social problems in the population, as well as the effect of changes made or proposed in social insurance and public assistance legislation upon the relative number and kinds of persons receiving public assistance in different parts of the country.

The coverage discussed thus far relates to individual programs. What is the net coverage of all the programs concerned with a given risk group? How many aged persons, for instance, are receiving income from Federal old-age and survivors insurance, the railroad and government retirement programs, the veterans program, or old-age assistance? It is, of course, possible to bring together the relevant figures, not only for the aged, but for the other risk groups as well, such as the survivors, the unemployed, and the disabled. The difficulty with these compilations is general lack of information on the number of persons receiving income from two or more programs. The size of the groups receiving both old-age and survivors insurance and old-age assistance, and both old-age and survivors insurance and aid to dependent children, has been available on a semi-annual basis since the middle of 1951, but with this exception, methods have still to be devised for eliminating duplication among beneficiary and recipient rolls. In their present form the data are useful in indicating the minimum size of the group at risk lacking protection.

Still another approach to net coverage, this time with reference to

State differences, is the use of per capita expenditures for aggregate social insurance benefits. State differences in this average presumably reflect differences in State provision for given risks (unemployment or temporary disability, for example); however, they also reflect program coverage limitations which have a differential impact on States because of differences in the industrial composition of the working population (old-age and survivors insurance, for example). State differences in the coverage of need may be measured, similarly, by reference to State per capital expenditures for public assistance. The latter average is affected, of course, by differences in the extent of need, as well as by differences in statutory provisions.

The raw materials for the appraisal of benefit adequacy in the social insurance and related programs include data on aggregate payments, the number of beneficiaries, the size distribution of benefit payments, the wage loss being compensated, the dependents of beneficiaries, the proportion of beneficiaries exhausting benefit rights (in the short-term programs), the income of beneficiaries, and living costs.

Aggregate payments are available for all programs; for the workmen's compensation and State and local government retirement programs, the aggregates are estimated. With the exception of workmen's compensation, data or estimates are also at hand on the number of beneficiaries. This means that for most programs the average benefit can be computed. Size distributions, which are a better guide to the analysis of benefit adequacy than the average, are published for the Federal old-age and survivors insurance system, the Railroad Retirement Board programs, the Federal Government's retirement system, and State unemployment insurance. Size distributions are not available, at least in published form, for other government retirement programs, the veterans program, the State temporary disability programs and workmen's compensation. All the short-run programs collect and tabulate data on the proportion of beneficiaries exhausting benefit rights.

The other types of data mentioned in the first paragraph of this section—the wage loss sustained by beneficiaries, the dependents of beneficiaries, the income of beneficiaries from benefits and other sources, and living costs for the most common family groupings encountered among beneficiaries, at selected living levels—are not, with some exceptions, the by-product of program operations and are obtainable only through special study. Especially noteworthy in this connection are the field surveys conducted by the Social Security Administration in the past 10 years of the income and living arrange-

ments of old-age and survivors insurance beneficiaries. A number of States have made similar studies in the field of unemployment insurance. With these exceptions there is a general lack of quantitative data in the social insurance programs on the relation of the benefit amount to family requirements and the place of the benefit in the family income structure.

Certain derived measures have been found useful in benefit appraisal which are not dependent upon field study. These include average benefit as a per cent of the average wage of all workers in covered employment (this is not the same thing as the average wage of beneficiaries prior to retirement, disability, or unemployment), maximum benefit under the law as a per cent of average wage in covered employment, per cent of benefits which are at the maximum, the per cent of beneficiaries receiving public assistance, and changes over time in the average benefit as compared with changes over time in the Consumer Price Index of the Bureau of Labor Statistics. To a greater or lesser extent measures of this kind have been developed in the Federal old-age and survivors insurance program and in the State unemployment insurance programs. They help answer such questions as, How much wage loss is compensated by the benefit? Has the benefit formula kept pace with changing wage and price levels? Has the insurance program reduced the need for public assistance? The answers to these questions have an important place in legislative consideration of the need for program amendment.

The tools for measuring the adequacy of payments in public assistance are, with some exceptions, those referred to in connection with the social insurance and related programs: data on aggregate payments, the number of recipients, the size distribution of payments, the income of public assistance recipients, the size and characteristics of public assistance families, the trend in average payment in relation to the trend in the cost of living, and general information on living costs in relation to family size and composition.

Statistical series on aggregate payments, recipients, average payment per recipient, and the size distribution of payments are available for all five types of public assistance on a monthly or less frequent basis. Average payment and size distribution measure differences among the States in the amount of assistance available to recipients. They reveal little, however, on the relation of the payment to requirements and on the extent of other income in the family. Special studies have developed information along these lines for old-age assistance and aid to dependent children recipients in selected States, but because

resources available to assistance recipients and their treatment by assistance agencies change with time, the applicability of the findings is limited to the period covered by the studies. Similar information is lacking for the other programs.

Major gaps in knowledge concerning the extent to which the income-maintenance programs achieve their objectives are summarized below.

With respect to coverage, there are no reported data and no regularly published estimates of the number of employees in employments covered by State and local government retirement systems or by workmen's compensation. Coverage and benefit data in three of the four State temporary disability insurance programs are incomplete. Current estimates of beneficiaries under workmen's compensation are lacking. With some exceptions, information is unavailable on the number of beneficiaries and public assistance recipients receiving income from two or more programs.

Benefit data are deficient in a number of respects. There is no information for the country as a whole on size of benefit in the fields of workmen's compensation and State and local government retirement systems. Average payments are available for veterans' benefits but not size distributions of benefits. Generally missing in all social insurance and related programs with the exception of Federal old-age and survivors insurance and some State unemployment insurance programs are any efforts at appraisal of the benefit in relation to other sources of income, family requirements, and living costs. With these exceptions, again, no materials have been developed to relate the benefit to the wage loss, or to estimate the number of beneficiaries receiving public assistance supplementation. In public assistance the chief deficiencies are in current data on the nonassistance income of recipients and organized and analyzed information on the budgetary standards used by public assistance agencies for estimating requirements.

With the exception of public assistance, unemployment insurance, and Federal old-age and survivors insurance, small-area data on beneficiaries and their benefits are generally lacking.

More difficult to classify because not directly related to program operations are deficiencies in the documentation of some of the commonest generalizations in the field of social security. We say the aged prefer insurance to assistance. How do we know this? Has anyone bothered to find out? To what extent does a contributory system of social insurance, with benefits paid as a matter of right and based on earnings, preserve "self-reliance and initiative . . . dignity and independence," reward "ambition and effort," and encourage private

savings, as claimed by the House Ways and Means Committee in 1949?[2] What proof do we have that the Bureau of Employment Security is correct in stating that unemployment insurance "sustains the morale and conserves the skills and standard of living" of the unemployed?[3] Is there any objective basis for the claim of the House Ways and Means Committee that under a contributory social insurance system "productivity is encouraged," the "Nation's total production is increased," and that such a system helps "to protect the Nation from serious economic maladjustment?"[2]

To ask these questions is to call attention to the complete absence of attitude research in the income-maintenance field. Such research may have no direct bearing on day-to-day program operations, but the lack of it is painfully apparent when public assistance comes under legislative or newspaper attack, or when the alleged merit or lack of merit of the reserve fund in old-age and survivors insurance is publicly aired. It would be helpful then to know what the general public thinks of the issues involved, what the beneficiaries or recipients think, and what the views are of the rank and file staff member working in the program, newspaper editors, legislators, and others.

Reference should be made also to the paucity of bench marks by which to measure benefit adequacy. The lacunae here are the primary responsibility of other kinds of programs. We lack detailed information on the relative importance of different sources of income at various income levels, at different ages, and among different family types and sizes. Research in this area would put in another light the social and political issues posed, for instance, by the contradiction between the income a good public assistance standard would provide and what the family earner in a low-paid job could earn.

NOTE

Published sources for the statistical series referred to in this article are listed below.

Earners, earnings, insurance status, benefit payments, and beneficiaries under social insurance and related programs:

Federal Security Agency, Social Security Administration, *Social Security Bulletin* (monthly), including Annual Statistical Supplement, published as part of the September issue.

Federal Security Agency, Social Security Administration, Bureau of Old-Age

[2] House of Representatives, Committee on Ways and Means, *Report to Accompany H.R. 6000*, Report No. 1300, 81st Congress, 1st session, August 22, 1949, pp. 2, 3.

[3] Department of Labor, Bureau of Employment Security, *Unemployment Insurance: Purposes and Principles*, Dec. 1950, p. 1.

and Survivors Insurance, *Quarterly Summary of Wage, Employment and Benefit Data* (Feb., May, Aug., Nov.).
—*Handbook of OASI Statistics* (annually).
Department of Labor, Bureau of Employment Security, *The Labor Market and Employment Security* (monthly), including monthly *Statistical Supplement* and quarterly supplement *Employment and Wages of Workers Covered by State Unemployment Insurance Laws by Industry and State.*
—*Significant Temporary Disability Data* (annually).
Railroad Retirement Board, *The Monthly Review.*
—*Annual Report.*
U. S. Civil Service Commission, *Retirement Report* (annually).
Veterans Administration, *Statistical Summary,* Claims (monthly).
—*Unemployment Allowance, Claims and Payment Data* (monthly).
—*Self Employment Claims and Payment Data, by Agencies* (Monthly).
—*Annual Report.*

All State unemployment insurance agencies issue recurrent data periodically, in greater or lesser detail. The publications of the California, New Jersey, and Rhode Island agencies contain, in addition, data on the temporary disability programs in these States.

Program data of varying degrees of completeness may be found in the reports of State workmen's compensation agencies. Some but by no means all State and local government retirement systems include beneficiary data in their annual reports.

Recipients and payments under public assistance programs:

Federal Security Agency, Social Security Administration, *Social Security Bulletin.* Recipients and payments appear monthly. Recipient rates are computed semi-annually and published in the March and October issues. Annual per capita expenditures for public assistance are to be found in the March issue.
—*Assistance Payments Under State-Federal Programs.* Annual release on the size distribution of payments for the month of September.

Periodic data are published by all State public assistance agencies.

Economic status of selected population groups:

Federal Security Agency, Social Security Administration, *Social Security Bulletin,* June and December. Semi-annual estimates of the number of aged persons, widows, and paternal orphans with income from employment, social insurance and related programs and public assistance.

General population and economic data useful in developing coverage estimates:

Department of Commerce, Bureau of the Census, *Monthly Report on the Labor Force* (*Current Population Reports, Labor Force, Series P-57*).
—Annual release, usually for the month of April, on marital status, family composition, and household characteristics of the population. (*Current Population Reports, Population Characteristics, Series P-20.*)
—Bureau of Foreign and Domestic Commerce, Office of Business Economics, *Survey of Current Business* (monthly), for estimates of personal income, by source, and State per capita income.

Some special studies referred to in the text:

Theodore D. Woolsey, "Estimates of Disabling Illness Prevalence in the United States." *Public Health Monograph No. 4*, Public Health Service, 1952.

Edna C. Wentworth, "Resources of Aged Insurance Beneficiaries: 1951 National Survey," *Social Security Bulletin*, August 1952.

Edna C. Wentworth, "Income of Old-Age and Survivors Insurance Beneficiaries, 1941 and 1949," *Social Security Bulletin*, May 1950.

Lelia M. Easson, "Adequacy of the Income of Beneficiaries Under Old-Age and Survivors Insurance," *Social Security Bulletin*, Feb. 1948.

The Role of Unemployment Compensation in Maintaining Family Income and Expenditure in an Area of Critical Unemployment. National Opinion Research Center, Chicago, 1951.

A Report on Benefit Financing and Solvency of the Employment Security Fund in Rhode Island. Rhode Island Department of Employment Security, Providence, 1950.

Aid to the Blind Recipients With Earnings in September 1950. Public Assistance Report No. 19, Social Security Administration, 1952.

Charles E. Hawkins, "Old-Age Assistance Recipients: Reasons for Nonentitlement to Old-Age and Survivors Insurance Benefits," *Social Security Bulletin*, July 1952.

Elizabeth Alling and Agnes Leisy, *Aid to Dependent Children in a Postwar Year.* Social Security Administration, 1950.

Ruth White and Thomas G. Hutton, *Requirements and Incomes of Recipients of Old-Age Assistance in 21 States in 1944*, Public Assistance Report No. 13, Social Security Administration, 1948.

"Living Arrangements and Physical Condition of Aged Recipients, 21 States," *Social Security Bulletin*, June 1947.

Ralph G. Hurlin, Sadie Saffian, Carl E. Rice, *Causes of Blindness Among Recipients of Aid to the Blind.* Social Security Administration, 1947.

Families Receiving Aid to Dependent Children, October 1942, Public Assistance Report No. 7, Social Security Administration, 1945.

AN APPRAISAL OF THE 1950 CENSUS INCOME DATA*

HERMAN P. MILLER
Bureau of the Census

THE establishment of a program designed specifically to measure the quality of the data obtained in the Census is one of the important differences between the 1950 Censuses of Population, Housing, and Agriculture and those previously conducted. Several reports on the quality of the data obtained in the Census, which will incorporate the findings of the Post Enumeration Survey (a reinterview survey conducted by highly trained enumerators after the Census was completed) and the record matching studies, are in various stages of completion. The purpose of the present article is to provide an interim report on the reliability of the 1950 Census income data.

In this paper, comparisons are made between the income size distributions obtained from preliminary tabulations of the 1950 Census and those obtained from the income survey conducted by the Census Bureau in March 1950 as a supplement to the Current Population Survey. In addition, estimates of aggregate income derived from each of these surveys are compared with independent estimates prepared by the National Income Division of the Department of Commerce.[1] The comparison of the aggregates emphasizes the well established fact that data in all field surveys of income are subject to errors of response and nonreporting which produce underestimates of the level of income.[2]

* The author is greatly indebted to Edwin D. Goldfield and Gertrude Bancroft of the Bureau of the Census who directed most of the work summarized in this paper. Thanks are also due to Leon Paley and Robert Stein for their able assistance in all phases of the work.

[1] All of the 1950 Census data to be discussed were obtained from a preliminary sample of Census returns which will be referred to as the Preliminary Sample Tabulations (PST). These data will be compared with information obtained in the Census Bureau's Current Population Survey (CPS) and with the income aggregates prepared by the National Income Division of the Department of Commerce (NID). Both the CPS and PST data are based on samples and are therefore subject to sampling variation. Figures based on relatively small numbers of cases, as well as small differences between figures, should be used with particular care. Estimates of the magnitude of sampling variation for the CPS figures may be obtained from the Census Bureau report, Series P-60, No. 7 and corresponding figures for the PST data may be found in the report, Series PC-7, No. 2.

[2] Virtually all analyses of income data obtained from household interviews show that this method produces an understatement of income. The most complete appraisal of the available data on the size distribution of income appears in Volume 13 of *Studies in Income and Wealth*, published by the National Bureau of Economic Research. The articles in this volume by Goldsmith; Mandel; and Wasson, Hurwitz, and Schweiger are particularly useful in this connection. The Bureau of Labor Statistics report, *Family Spending and Saving in Wartime*, Bulletin No. 822, contains a detailed appraisal of the income data obtained in the Survey of Family Spending and Saving in Wartime as well as brief indications of the extent of underreporting of income in other field surveys such as the Consumer Purchases Study, 1935–1936; Minnesota Income Study, 1938–1939; and the 1940 Census. The only detailed appraisal of the wage or salary income data obtained in the 1940 Census appears in an unpublished memorandum prepared by the staff of the Census Bureau.

The Census Bureau has noted in almost all of its income reports that income information obtained in household interviews is largely "based on memory rather than on records, and in the majority of instances on the memory or knowledge of some one person, usually the wife of the family head. The memory factor in data derived from field surveys of income probably produces underestimates because the tendency is to forget minor or irregular sources of income."[3] At the same time, the data suggest that despite these deficiencies, the information obtained in the field surveys may be sufficiently reliable for most uses.

In addition to the defects inherent in the household interview approach, the 1950 Census had some special problems in regard to the collection of income data. For example, the enumerators used in the Census were less skillful, less adequately prepared for their jobs, and more burdened with procedural details than are the enumerators used in the Current Population Survey (CPS). In addition, although the "line" schedule used in the 1950 Census was excellent for certain purposes, it was not well adapted for the collection of family income data.[4] Despite these handicaps, the 1950 Census income data show distributions quite similar to those obtained in the CPS, which has produced a consistent and useful series of national estimates of the distribution of consumer income for each year since 1944. The similarity of the CPS and Census income data despite the limitations noted above may in part merely reflect the fact that within the framework of the collection techniques currently employed by the Census Bureau, the differences which will be obtained by alternative forms of questioning will not be striking. This was demonstrated in one of the 1950 Census pretests.[5] It is also possible that some of the defects inherent in the 1950

[3] U. S. Bureau of the Census, *Current Population Reports—Consumer Income*, Series P-60, No. 9, "Income of Families and Persons in the United States: 1950," March 25, 1952, p. 19.

[4] The family income data in the 1950 Census were obtained by asking questions separately (a) for the head of the family, and (b) for all relatives of the head *as a group*. This method does not probe as deeply as the CPS income supplements in which income questions are asked individually for each member of the family.

[5] Several different types of income questions were tested by the Bureau of the Census in the April 1948 Current Population Survey for the purpose of developing a simplified method of obtaining income data, suitable for the 1950 Census. A total of 25,000 households were interviewed by trained enumerators and the following alternative sets of income questions were asked of systematic subsamples of the total sample.

Type 1.—On one-half of the schedules the enumerator was to ask the specific amount of (a) wages and salaries; (b) total money income; for each person 14 years and over.

Type 2.—On one-fourth of the schedules the enumerator was to ask the specific amount of total money income for each primary family and for each person not a member of a primary family.

Type 3.—On one-fourth of the schedules the enumerator was to ask the respondent to indicate on a card the broad class interval into which the total money income fell for each primary family and for each person not a member of a primary family.

The findings of the field tests of each of the questions are presented in the table below. One of the

Census were balanced by better respondent cooperation on the average, despite occasional strong individual objection. The legal provisions of the Census may have stimulated more accurate reporting on the part of respondents.

TABLE 1

PST AND CPS ESTIMATES OF NUMBER OF INCOME RECIPIENTS AND MEDIAN INCOME, BY TYPE OF INCOME, FOR THE UNITED STATES: 1949

Number of income recipients and median income, by type of income	PST (Total population 14 years old and over)	CPS (Noninstitutional population 14 years old and over, excluding members of armed forces on post)
Total population	111,926,000	109,644,000
Total number of income recipients	72,054,000	71,768,000
Number of recipients of:		
Wage or salary income	54,974,000	54,912,000
Self-employment income	10,674,000	10,966,000
Income other than earnings	18,813,000	16,875,000
Median income:		
Total income	$1,909	$1,814
Wage or salary income	$2,032	$2,016
Self-employment income	$1,586	$1,039
Income other than earnings	$ 478	$ 496

INCOME OF PERSONS

1. *Comparison of CPS and PST income data for persons.* Essentially the same techniques were used to obtain personal income data in the March 1950 Current Population Survey and in the 1950 Census. Therefore, it is of some interest to compare the results. Table 1 presents a summary of the numbers of income recipients and medians for each

striking conclusions derived from these tests was that asking for a global figure on total income for the family *group* (type 2) and asking for information on wages or salary and total income for *each* family member (type 1) produced about the same median income.

MEDIAN TOTAL MONEY INCOME FOR FAMILIES AND INDIVIDUALS BY TYPE OF SCHEDULE, FOR THE UNITED STATES, URBAN AND RURAL: 1947

Area	Type 1	Type 2	Type 3
United States	$2,702	$2,741	$2,581
Urban	2,961	2,985	2,741
Rural nonfarm	2,607	2,620	2,544
Rural farm	1,760	1,896	1,923

type of income. Although the data require adjustment to make them comparable with respect to population coverage (see text below), they provide a basis for comparing the PST and CPS income distributions at the most crucial points.

With the exception of income other than earnings, PST and CPS appear to have reported about the same number of recipients of each type of income. The apparent similarity of these estimates is somewhat misleading inasmuch as the two surveys are not directly comparable with respect to population coverage. The PST included about 1.6 million institutional inmates,[6] and about 0.3 million members of the armed forces living on military posts who were excluded from CPS. Since these persons are in the population base from which the PST estimates of the numbers of income recipients are made, it is apparent that the PST estimates are somewhat overstated as compared with CPS.[7] If the institutional inmates and members of the armed forces living on post were excluded from the PST weighting of the data, the total number of income recipients in PST would have been estimated to be about 70.6 million rather than 72.1 million as shown in the table. Experience in the reconciliation of other estimates of this type indicates that the discrepancy of about 1 million between the CPS and the adjusted PST estimate of the number of income recipients is relatively small. For example, considerably greater differences in the estimates of the number of paid workers during a given year are usually obtained from the CPS surveys of work experience conducted in December and the income surveys conducted in April.

It is also apparent from Table 1 that with the exception of income from self-employment, the medians obtained for each type of income in PST and CPS were about the same. However, a closer look at the full distributions by income levels in Table 2 indicates that PST consistently shows a larger proportion of income recipients in the upper-income brackets. The percentage differences do not appear great; however, since the upper-income groups possess a large share of the aggregate income, the small percentage differences shown in the table yield relatively large differences in the aggregates. (See page 33 ff. for a detailed discussion of the aggregate comparisons.)

[6] Income information was very poorly reported for institutional inmates in the 1950 Census. Of the 1.6 million inmates of institutions 14 years of age and over, an estimated 1.1 million did not report on income, 0.3 million reported no income, and 0.2 million reported $1 or more of income.

[7] No adjustment was made for the fact that the PST estimate of the civilian noninstitutional population is 0.4 million greater than the CPS estimate since this represents a difference in the independent estimate of the same population rather than a difference in population coverage. The difference between these two figures arises from the fact that the population estimate used to inflate the CPS data is a projection of 1940 Census data, whereas the PST estimate represents an inflation of a sample of 1950 Census returns.

The discrepancy of approximately $600 between the PST and CPS estimates of median self-employment income is rather large. A full explanation for this difference is not available at present. Perhaps the Census obtained more accurate income data than CPS in farm areas because of the use of the detailed farm schedule in the Census of Agriculture, which tended to improve the estimate of net income from farm self-employment on the population schedule. This improvement would have been reflected in less underreporting of income or in the higher

TABLE 2

PST AND CPS DISTRIBUTIONS OF PERSONS 14 YEARS OF AGE AND OVER AND OF AGGREGATE MONEY INCOME, BY TOTAL MONEY INCOME, FOR THE UNITED STATES: 1949

Total money income	PST		CPS	
	Total population 14 years old and over	Aggregate income	Noninstitutional population 14 years old and over (excluding members of armed forces on post)	Aggregate income
	(thousands)	(billions)	(thousands)	(billions)
Total	111,926	$171.4	109,644	$158.0
Total with income	72,054	—	71,768	—
Per cent of those with income.	100.0	100.0	100.0	100.0
Loss or $1 to $499	17.3	1.8	18.9	2.1
$ 500 to $ 999	13.5	4.3	13.8	4.7
$ 1,000 to $1,499	10.9	5.7	10.8	6.1
$ 1,500 to $1,999	10.2	7.5	10.4	8.3
$ 2,000 to $2,499	11.2	10.6	11.6	11.9
$ 2,500 to $2,999	9.1	10.6	9.5	11.9
$ 3,000 to $3,499	9.0	12.3	8.7	12.8
$ 3,500 to $3,999	5.6	8.9	5.4	9.2
$ 4,000 to $4,499	4.0	7.1	3.5	6.8
$ 4,500 to $4,999	2.2	4.4	2.0	4.3
$ 5,000 to $5,999	2.9	6.8	2.3	5.7
$ 6,000 to $6,999	1.3	3.6	1.1	3.2
$ 7,000 to $9,999	1.4	5.0	1.0	3.9
$ 10,000 and over	1.4	11.5	1.0	9.1
Median income	$1,909	—	$1,814	—

median income of farm families. (See Table 6.) It is also possible that the Census income data benefited from the self-enumeration procedure which was supposed to have been used in the Census of Agriculture in all farm areas except the South; however, in the absence of any firm data on the effects of self-enumeration or the extent to which it was actually used in the Census of Agriculture, this point can only be a matter of conjecture. The only available evidence on this point comes from a pretest of the Census conducted in October 1948 in Union County, Indiana, and Carroll County, Kentucky, to determine the relative merits of self-enumeration and direct-enumeration techniques. This pretest showed that in about 60 per cent of the cases where the self-enumeration procedure should have been used, the Census enumerators rather than the respondents actually completed the schedules. Even where the respondents entered the income information on the self-enumeration schedules, a large proportion of the entries were apparently rounded to hundreds of dollars indicating that the respondents probably did not take advantage of the advance distribution of schedules to refer to records. Moreover, there were no significant differences in the medians produced by each procedure.

2. *Comparison of CPS, PST, and NID aggregates.* Although the primary purpose of the income questions in the 1950 Census and in the March 1950 Current Population Survey was to provide a distribution of the population by income levels, estimates of the aggregates of each type of income can be derived from these data. The comparison of these estimates with those prepared by the National Income Division of the Department of Commerce provides further information regarding the reliability of the 1950 Census income data. Such a comparison is made in Table 3. The Census aggregates shown in this table largely represent estimates derived from the tabulated data and are subject to substantial errors of estimation. Moreover, despite the fact that the NID aggregates are largely based on record data and may therefore be considered more reliable than the field survey data, they are also subject to errors of estimation. For these reasons, small differences in the results must be interpreted with care.

The PST and CPS estimates of aggregate income shown in the table below were computed by multiplying the estimated number of persons in each income interval by the average income for that interval. For intervals between $500 and $10,000, the midpoint of each interval was assumed to be the average; $250 was selected as the average for the "Loss or $1 to $499" level, and $20,000 was selected as the average

for the $10,000 and over level.[8] The aggregates were obtained by multiplying the frequencies by the average income for each interval. The NID aggregates with which the PST and CPS aggregates are compared were obtained by adjusting the Personal Income Series of the Department of Commerce to make it more comparable with the definition of income and the population coverage of the Census income data.[9]

TABLE 3

PST, CPS, AND NID ESTIMATES OF AGGREGATE INCOME, BY TYPE OF INCOME, FOR THE UNITED STATES: 1949

(Billions)

Type of income	PST (Total population 14 years old and over)	CPS (Noninstitutional population 14 years old and over, excluding members of the armed forces on post)	NID (Total population 14 years old and over)
Total income	$171.4*	$158.0*	$187.6
Wage or salary income	124.3	120.0	127.5‡
Self-employment income	31.1	26.5†	29.3
Income other than earnings	16.6	13.3	30.8

* The detail by type of income does not add to the total because each of the aggregates was estimated independently.

† Nonfarm self-employment estimate: $19.1 billion; farm self-employment estimate: $7.4 billion.

‡ Nonfarm self-employment estimate: $19.4 billion; farm self-employment estimate: $9.9 billion.

Perhaps the most striking fact shown in the above table is that the PST obtained about 90 per cent of the comparable NID estimate of aggregate income. This proportion is considerably greater than that obtained in any of the CPS income surveys which have been conducted to date. The March 1950 CPS obtained only 84 per cent of the NID aggregate.

[8] The arithmetic mean for the "tail" of an income distribution having an open-end interval of $10,000 or more is generally estimated from a Pareto curve. However, it was found after some investigation that although this method yields reasonable estimates for the distributions of families and unrelated individuals, it tends to overstate the mean for distributions of persons. An examination of income tax returns for 1944–1949 showed that in each of these years the mean adjusted gross income for returns with $10,000 or more of adjusted gross income was between $21,000 and $22,000. On the basis of this and other evidence it was assumed in deriving the data shown in Tables 2 and 3 that $20,000 was a reasonable approximation of the mean for the open-end interval.

[9] A detailed description of the method used to adjust the Personal Income Series for comparability with the Census income data is described in a paper by S. F. Goldsmith, "Appraisal of Basic Data Available for Constructing Income Size Distributions," which appears in *Studies in Income and Wealth*, Volume 13 (National Bureau of Economic Research).

For the reasons noted in a preceding section, the PST and CPS estimates of aggregate income shown in the table are not directly comparable. If institutional inmates are excluded from the PST tabulations, the estimated aggregate income is reduced from $171.4 billion to $168.0 billion. An additional adjustment to exclude members of the armed forces living on post from the PST data reduces the aggregate about $1 billion more to about $167.0 billion. Thus, the actual difference between the PST and CPS aggregates when adjusted for population coverage is about $9 billion.

TABLE 4

PERSONS 14 YEARS OF AGE AND OVER BY TOTAL MONEY INCOME IN CPS AND IN THE CENSUS, FOR THE UNITED STATES

Comparison of CPS and Census	All persons	Paid workers		
		Total	Wage or salary workers	Self-employed workers
Total in sample	5,701	3,118	2,529	589
Total reporting on income in CPS and Census	4,898	2,607	2,165	442
Per cent in same income interval in both*	61	49	52	31
Per cent in higher income interval in CPS	20	27	25	38
Per cent in higher income interval in Census	18	25	23	31
Median total money income				
CPS	$1,914†	$2,326	$2,389	$1,797
Census	$1,849†	$2,238	$2,319	$1,543

* The income intervals used were those shown in Table 2 with the exception that $6,000 to $9,999 was treated as a single interval.

† These figures differ from those shown in Table 2 because of sampling variation.

3. *Income data from CPS-Census Match.* About one-fifth of the persons who reported on income in the March 1950 CPS were also required to furnish income data in the 1950 Census. The differences between the two sets of reports show the extent of variability of response in these surveys. They also indicate some of the concomitant effects of this variability on the income distributions. The data in Table 4 are based on an analysis of the income reported by 5,700 persons 14 years of age and over who were interviewed in both CPS and in the 1950 Census.

Of the 4,898 persons who reported on income in both surveys, 61 per cent were in the same income interval in both surveys, 20 per cent were in a higher income interval in CPS, and 18 per cent were in a higher income interval in the Census. (About 82 per cent of the persons reporting on income in both surveys were either in the same interval or in one higher or lower adjacent interval.) Despite the fact that nearly two-fifths of the respondents were reported at different income levels in each survey, there was no significant difference between the medians or between the distributions by income levels. The similarity in the distributions largely reflects the fact that income was relatively overstated in each survey about as frequently as it was understated. These reporting and enumerative errors generally tended to cancel, leaving the income distributions relatively unchanged. Despite the similarity of the over-all distributions, there may be important differences between the CPS and Census income distributions for certain segments of the population. The relatively small size of the CPS-Census Match sample precludes any detailed examination of the income effects associated with variability of response at the present time. A detailed study of this characteristic will be included in the analysis of the results of the Post Enumeration Survey. However, the analysis of income differences by class of worker (Table 4) and by age, sex, and veteran status (Table 5) does not indicate that variability of response resulted in any significant differences for these characteristics.

Over four-fifths of the persons reporting no income in the March 1950 CPS also reported no income for 1949 in the Census, but only about one-half of the persons reporting some income in CPS were in the same income interval in the Census. Some indication of the extent of variation in response among income recipients may be obtained from the figures for paid workers in Table 4. About one-half of the wage or salary workers were in different income intervals in both surveys; however, the medians for these workers do not differ significantly. As might be expected, there was considerably greater variation of response for self-employed workers than there was for wage or salary workers. About 70 per cent of these workers reported different incomes in the two surveys. Even here, however, incomes were underreported almost as frequently as they were overreported.

Table 5 indicates that about three-fourths of the women as compared with only one-half of the men reported the same income in both surveys. However, nearly half of the women in the sample but only 6 per cent of the men reported no income. Among income recipients,

TABLE 5

CONSISTENCY OF INCOME REPORTING IN CPS AND IN THE CENSUS FOR PERSONS 14 YEARS OF AGE AND OVER BY SEX, AGE, AND VETERAN STATUS

Sex, age, and veteran status	Per cent in same income interval in CPS and the Census		Median income	
	Persons reporting on income in CPS and Census	Income recipients in CPS and Census	CPS	Census
Total	61	44	$1,914	$1,849
Male	49	44	2,432	2,378
14 to 24, Total	56	43	1,054	1,289
14 to 24, Veteran of World War II	48	48	2,185	2,231
14 to 24, Nonveteran of World War II	58	41	722	897
25 to 44, Total	43	43	2,904	2,850
25 to 44, Veteran of World War II	43	43	2,955	2,918
25 to 44, Nonveteran of World War II	43	43	2,860	2,773
45 to 64	46	46	2,644	2,646
65 and over	48	44	1,037	1,234
Female	73	43	910	941
14 to 24	72	47	838	992
25 to 44	73	43	1,271	1,293
45 to 64	72	41	888	915
65 and over	69	45	575	497

about the same proportion (one-half) were in the same income level for both sex groups. The medians and distributions by income level in CPS and the Census differed by negligible amounts for both males and females.

Within each sex group, the proportion of income recipients who were in the same income level in both surveys did not vary by age or veteran status. The differences by age and veteran status in the median incomes derived from the two surveys were also insignificant.

INCOME OF FAMILIES

1. *Comparison of CPS and PST income distributions.* The family income data obtained in the March 1950 CPS represent about the best

estimate of the distribution of families by income level which can be made within the limitations of the collection techniques thus far used by the Census Bureau. For this reason, despite their shortcomings, these data provide a standard with which the 1950 Census income data can be compared.

Table 6 shows that there is no significant difference between the PST and CPS median income for families within each color and residence group.[10] The more detailed distributions by income level and color in Table 7 indicate a similar correspondence between the results obtained in these surveys. The one marked difference between the two distributions is the larger proportion of "zero-income" families obtained in the PST results. This deficiency in the PST data resulted largely from the enumerative and editing procedures used in the 1950 Census and is discussed in some detail below.

TABLE 6

PST AND CPS ESTIMATES OF MEDIAN TOTAL MONEY INCOME OF FAMILIES, FOR THE UNITED STATES, URBAN AND RURAL

Area	All classes		White		Nonwhite	
	PST	CPS	PST	CPS	PST	CPS
United States.....	$3,068	$3,107	$3,216	$3,232	$1,425	$1,650
Urban.....	3,429	3,486	3,581	3,619	1,850	2,084
Rural nonfarm.....	2,552	2,763	2,711	2,851	1,141	1,240
Rural farm.....	1,734	1,587	1,936	1,757	733	691

Although the medians obtained in CPS and PST are substantially the same, patterns of small differences which appear in the data may be of some significance. Among both white and nonwhite families the CPS medians in urban and rural-nonfarm areas are slightly (though not significantly) higher than the corresponding PST results; however, among rural-farm families PST showed somewhat higher medians than CPS. These differences may reflect differences in the procedures used in these areas. In nonfarm areas essentially the same procedures were used in CPS and Census. Although a small part of the differences between CPS and PST are attributable to the fact that college students living away from home were included as family members in CPS but not in the Census, it is likely that the higher CPS results in these areas

[10] Unrelated individuals are excluded from this comparison because they are defined somewhat differently in CPS and in the Census. For example, college students living away from home are treated as family members in CPS and as unrelated individuals in the Census.

reflect the superior schedule design and better enumeration. The reasons for the higher PST estimates of farm income have already been discussed (see page 32).

Table 7 indicates that there were perhaps four times as many families without income in PST as in CPS. In absolute terms these families numbered 1.6 million in PST as compared with 0.4 million in CPS. The relatively large number of families without income in PST is partly accounted for by the editing procedures used in the 1950 Census. About one-third, or 0.5 million families without income in PST, were cases in which the family head had no income and income information for other relatives was not obtained. Although many of these families probably represented cases in which the head had no income but other family members did, they were all tabulated as having no income according to the income editing rules used in the 1950 Census (see discussion of

TABLE 7

PST AND CPS DISTRIBUTIONS OF FAMILIES BY TOTAL MONEY INCOME, BY COLOR, FOR THE UNITED STATES: 1949

Total money income	Total		White		Nonwhite	
	PST	CPS	PST	CPS	PST	CPS
Number (thousands)	38,788	39,193	35,411	35,988	3,377	3,205
Total	100.0	100.0	100.0	100.0	100.0	100.0
Percent reporting	93.9	*	93.8	*	95.9	*
Per cent not reporting	6.1	*	6.2	*	4.1	*
Total reporting	100.0	100.0	100.0	100.0	100.0	100.0
None	4.1	0.9	3.9	0.9	6.3	0.9
Loss, or $1 to $499	4.5	5.0	3.7	4.2	11.9	14.0
$ 500 to $ 999	6.8	6.2	5.7	5.3	18.4	16.0
$ 1,000 to $1,499	7.3	7.3	6.5	6.6	15.8	15.1
$ 1,500 to $1,999	7.5	7.6	7.0	7.1	12.6	13.5
$ 2,000 to $2,499	9.3	10.2	9.0	10.0	11.9	12.9
$ 2,500 to $2,999	9.0	10.4	9.2	10.5	7.0	9.2
$ 3,000 to $3,499	10.9	11.2	11.4	11.7	6.2	5.6
$ 3,500 to $3,999	8.7	8.8	9.3	9.3	3.0	3.7
$ 4,000 to $4,499	7.2	6.7	7.7	7.1	2.1	2.6
$ 4,500 to $4,999	4.9	5.3	5.3	5.6	1.0	1.9
$ 5,000 to $5,999	7.9	7.8	8.4	8.3	1.8	2.5
$ 6,000 to $6,999	4.3	4.8	4.6	5.1	1.0	1.3
$ 7,000 to $9,999	4.7	5.0	5.1	5.4	0.8	0.5
$10,000 and over	2.9	2.6	3.2	2.8	0.2	0.3
Median income	$3,068	$3,107	$3,216	$3,232	$1,425	$1,650

* Comparable figures not available.

effects of Census editing procedures below). At present little is known about the remaining 1.1 million "zero-income" families. Nearly half of these families probably had no income as defined in the Census, if the CPS figure of 0.4 million is taken to be accurate. They may have been living on savings, charity, or gifts, or else they were newly created families or families in which the sole breadwinner had only recently died or left the family. The remainder of these families appear at the "zero-income" level for a variety of reasons, such as errors in enumeration and coding.

2. *Comparison of PST aggregates obtained from the tabulations for families and unrelated individuals and for persons.* The PST tabulations for families and unrelated individuals yielded a considerably lower aggregate ($154.5 billion)[11] than the comparable tabulations for persons 14 years old and over ($168.0 billion). Theoretically, the two estimates should have been about the same. The lower estimates derived from

FOR PERSONS 14 YEARS OF AGE AND OVER

Income received by this person in 1949				If this person is a family head (see definition below) Income received by his relatives in this household			
Last year (1949), how much money did he earn working as an employee for wages or salary? (Enter amount before deductions for taxes, etc.)	Last year, how much money did he earn working in his own business, professional practice, or farm? (Enter net income)	Last year, how much money did he receive from interest, dividends, veteran's allowances, pensions, rents, or other income (aside from earnings)?	LEAVE BLANK	Last year (1949), how much money did his relatives in this household earn working for wages or salary? (Amount before deductions for taxes, etc.)	Last year, how much money did his relatives in this household earn in own business, professional practice, or farm? (Net income)	Last year, how much money did his relatives in this household receive from interest, dividends, veteran's allowances, pensions, rents, or other income (aside from earnings)?	LEAVE BLANK
31a	31b	31c	F	32a	32b	32c	G
☐ None $_____	☐ None $_____	☐ None $_____		☐ None $_____	☐ None $_____	☐ None $_____	
☐ None $_____	☐ None $_____	☐ None $_____		☐ None $_____	☐ None $_____	☐ None $_____	

[11] The average for the open-end interval was estimated from a Pareto curve in the tabulations for families and unrelated individuals, whereas $20,000 was assumed as the average for the open-end interval in the tabulations for persons. If $20,000 had been used as the average in the family tabulations, the aggregate would have been $159.1 billion rather than $154.5 billion.

the tabulations for families are largely attributable to the method used to collect family income data in the Census.

Income information in the 1950 Census was obtained from a 20-per cent sample of the population. If a person came into the sample, he was asked questions 31a–c indicated in the excerpt from the 1950 Census schedule shown below. If the sample person was not a family head, questions 32a–c were skipped; however, if he was a family head, questions 32a–c were asked in order to obtain income information for the entire family group. Although this procedure provided an unbiased sample of families and of persons, it increased the possibility of introducing reporting errors.

TABLE 8

PER CENT OF FAMILIES AT EACH INCOME LEVEL UNDER $10,000 WITH FAMILY INCOME GREATER THAN HEAD'S INCOME: PST AND CPS*

Total money income	PST	CPS
Under $1,000	19.3	27.1
$1,000 to $1,999	30.7	40.6
$2,000 to $2,999	29.6	38.0
$3,000 to $3,999	33.9	41.3
$4,000 to $4,999	49.1	56.0
$5,000 to $5,999	60.0	70.1
$6,000 to $6,999	64.5	72.4
$7,000 to $9,999	65.6	74.0

* PST data showing the number of families with more than one income recipient for the $10,000 and over interval not available.

As the collection procedure implies, the family income data are based on replies to income questions asked separately for the head of the family and for all relatives of the family head *as a group*, whereas the income data for persons are based on replies to income questions asked for *each* person in the sample. The failure to obtain income information individually for relatives of family heads probably resulted in underreporting of income for this group. Evidence supporting this view is shown in Table 8 where it may be noted that at each income level the CPS data have a larger proportion of families in which the family income was greater than the head's income. This is equivalent to saying that there were more families having income recipients other than the head in CPS than in PST.

3. *Effects of Census editing procedures.* In the interest of economy, a relatively simple editing procedure was used in processing the 1950 Census family income data. It was assumed that if the income informa-

tion was obtained for the family head, the total family income was known even when the income information was not obtained for other family members. This procedure is different from the CPS procedure where any incompleteness in the returns for a family would result in a not reported (NA) classification for the family. It was adopted be-

TABLE 9

PER CENT DISTRIBUTION OF FAMILIES BY TOTAL MONEY INCOME, USING DIFFERENT NOT REPORTING (NA) CRITERIA

Total money income	Distribution as tabulated	Distribution if families in which income information was not obtained for relatives of the head were tabulated as NA
Total	100.0	100.0
Reporting	93.9	88.5
Not reporting	6.1	11.5
Total reporting	100.0	100.0
None	4.0	2.8
Loss or $1 to $499	4.6	4.3
$ 500 to $ 999	6.8	6.7
$ 1,000 to $1,499	7.3	7.3
$ 1,500 to $1,999	7.5	7.4
$ 2,000 to $2,499	9.3	9.3
$ 2,500 to $2,999	9.1	9.2
$ 3,000 to $3,499	11.0	11.1
$ 3,500 to $3,999	8.7	8.9
$ 4,000 to $4,499	7.2	7.4
$ 4,500 to $4,999	4.9	5.0
$ 5,000 to $5,999	7.9	8.1
$ 6,000 to $6,999	4.3	4.4
$ 7,000 to $9,999	4.7	4.9
$10,000 and over	2.9	3.0
Median income	$3,069	$3,135

cause it was one of the few editing procedures which would permit fuller utilization of all information obtained in the Census. Some of the pretests indicated that a relatively large proportion of the schedules submitted by the Census enumerators would contain blanks in the items for relatives of the head which actually represented the absence of income rather than the failure to report the amount of income received. Another consideration which led to the adoption of this procedure is the fact that in about one-half of all families the head is the only income recipient and even where there are other recipients the

head is generally the principal recipient. Although other editing procedures were considered, only the one finally adopted fitted in as part of the general coding and editing scheme used in the Census.

In general, the editing scheme used in the 1950 Census produced the desired results of keeping the NA rate low without seriously distorting the income distribution (see Table 9). The NA rate obtained using the 1950 Census editing procedures was half of what it would have been if the alternative procedure had been used of classifying as NA families in which no report was made for family members other than the head. The median income based on the procedure which was used was about $70 lower than the median which would have been obtained from the alternative procedure. Nevertheless, this procedure created a downward bias in the statistics inasmuch as all the missing entries were converted to zeros, although some of them must have actually represented amounts. The approximate magnitude of this bias is reflected in the fact that the aggregate computed from the distribution of families and individuals by income levels prior to editing was $4 billion higher than the aggregate derived from the data as tabulated (after editing). This suggests that editing accounts for about one-fourth of the difference between the family aggregate and the aggregate for persons.

SUMMARY

The 1950 Census data which are currently becoming available provide, for the first time, statistics on the size distribution of income for each locality in the United States, as well as analytical tabulations for the country as a whole which could not be obtained from the sample surveys conducted in earlier years. Experienced users of Census data have learned to inquire very carefully into the quality of information obtained by a mass enumeration of the population. For this reason, an investigation of the quality of these data is appropriate at the present time. Unfortunately, a full report on the quality of the income data will not be available until the results of the Post Enumeration Survey and the various record matching projects are completed. However, the preliminary investigation described above suggests that the 1950 Census income data should be satisfactory for most purposes. The two major conclusions of the investigation are: (1) the family income data obtained in the Census are of comparable level with those obtained in the annual income supplements to the Current Population Survey; and (2) the income data for persons obtained in the Census are somewhat more reliable than the family income data largely because the latter were obtained by an inferior collection technique necessitated by the use of a line schedule in the 1950 Census.

DELIMITATION OF ECONOMIC AREAS: STATISTICAL CONCEPTIONS IN THE STUDY OF THE SPATIAL STRUCTURE OF AN ECONOMIC SYSTEM*

RUTLEDGE VINING
University of Virginia

A RECENT monograph by D. J. Bogue describes the procedures used by the Bureau of the Census in grouping the counties of the United States into a new set of 501 areas called *State Economic Areas.* Along with the description of procedure, there is a large map showing the Areas, a table listing the counties in each Area, a table giving for each Area some 88 pieces of numerical data regarded as indicating the non-agricultural characteristics of the Area, and a similar table giving for each Area about 75 pieces of numerical data presented as indicating the agricultural characteristics of the Area.

The present discussion will concentrate upon the nature of these Areas that are designated as "functional groupings of counties" each containing a "distinctive economy." Interpreted simply as a new set of areas for Census reporting intermediate in size between the county and the State, the procedure described makes quite good sense. But there are overtones in the discussion that will suggest to some readers that more was intended, that the work was carried on to such effect that the component "economies" or "regional economic units" of this nation have been discovered and approximately designated. It is this that many readers will look for in such a work, and I believe this seeking for a "natural" area unit for economic studies is based upon a fundamental misunderstanding.

In my opinion, the spatial structure of a human economy should be regarded conceptually as virtually a continuum. As in other studies of phenomena having volume or spatial extension, empirical observations must be made upon the contents of finite and arbitrary spatial units, these empirical observations being viewed as providing an approximate conception of what would be viewed were the spatial units made arbitrarily smaller while the contents were being made more dense. From this point of view, I believe the student of the structure, functioning, and development of an economic system may regard as not particularly relevant the criteria used by the Bureau of the Census in

* Review article on Donald J. Bogue, *State Economic Areas—A description of the procedure used in making a functional grouping of the counties of the United States,* Bureau of the Census, Washington, D. C., 1951. This paper was done with the assistance and cooperation of the Bureau of Population and Economic Research, University of Virginia.

the designation of the Areas. While census data classified for 501 areas provide more information than if classified for only 48 States, less information is provided in this way than if they were classified by the 3070 counties. I should be inclined to interpret the choice of the number 501 as an economic decision in the sense that presumably the cost of making this number 502 was regarded by some responsible person or group as being greater than the value of the information that would be added; and the particular boundaries selected I believe should and can be defended on grounds other than those that may be found in the needs of political economists studying how a human economy works and develops as an organization.

I shall attempt to develop these points, first outlining the nature of the purposes that seem to me to give meaning to the criteria described as having been used in the designation of the State Economic Areas. In the monograph, these purposes are confused, I think, with others for which the criteria are not relevant. Accordingly, I shall discuss these other purposes in outlining a conception of regional or spatial structure in the specification of which such areas as are described in this monograph would play no essential role.

I

Those who were responsible for delimiting the Areas were confronted with the problem of selecting a working rule which, together with statistical observations, would determine the boundaries of sub-areas into which the entire territory of the nation is to be partitioned. An indefinitely large number of rules would of course be possible alternatives to the one finally selected. The selection of a particular rule would be made on the basis of how that rule performs in designating the county groupings, this performance to be measured in terms of the usefulness of the county groupings so designated. The technical specialists, who did the staff work and who professionally advised those making the final choice, had the problem of analyzing the respective performance characteristics of the many possible rules. The decision-makers presumably chose that rule which they regarded as having the most desirable performance properties.

The question that I raise now is this: In choosing a "best" rule for delimiting the areas, what purposes did the decision-makers have in mind for these particular area designations to serve? The rule selected would possess optimal properties in its capacity for least expensively generating a set of areas that attains predetermined objectives. What were these objectives?

The primary responsibility of the Bureau of the Census is that of providing enumerated information. It gauges the current demands for information, anticipates future demands, recommends to a "higher" authority budget expenditures upon information-gathering, seeks out economical means for obtaining specified information, helps this "higher" authority compare the costs of obtaining an increment to the total information gathered against an evaluation of the worth of this increment. It is continually being bombarded with requests for factual data from public and private administrative organizations. The problem always confronting the Bureau is that of choosing best courses of action in meeting the demands made upon it for information from the thousands of individual administrative decision-making units.

From this point of view, the monograph makes good sense in describing the procedure adopted by the Bureau in defining its new geographic classification:

> In tabulating and publishing data, the Bureau of the Census has found an increasing need for a set of areas intermediate in size between counties and States. For many statistical purposes, State units are too large and heterogeneous in their composition; whereas in many other instances county units are too small and too numerous to be usable. In the 1950 Census of Agriculture, for example, the use of State economic areas will provide cross-tabulations in considerable detail of data that previously have been available only at the State level. If made for individual counties, such tabulations would be costly and more detailed than needed for most purposes. . . . State economic areas are relatively homogeneous subdivisions of States. They consist of single counties or groups of counties which have similar economic and social characteristics. The boundaries of these areas have been drawn in such a way that each State is subdivided into a few parts, with each part having certain significant characteristics which distinguish it from the other areas which it adjoins. . . . In general, wherever it is not imperative that totals be reported for each county, the State economic areas may be used to present a concise body of statistics for the entire Nation, by States and their principal parts. Considerable savings in publication space, tabulation costs, and clerical work can be made through the use of these units instead of counties. Also, State economic areas permit the tabulation of sample statistics in much more detail than is possible for individual counties. It is for these reasons that State economic areas have been established by the Bureau. (p. 1)

The rule selected by the Bureau for delimiting the areas was substantially as follows:

1. The delimitation was made on the basis of statistical and other objective evidence; and homogeneity with respect to certain prescribed economic and social indices was "the prime criterion in judging the quality of the State economic areas delimitation."
2. Each area was required to satisfy minimum size requirements.

Those areas intended for general tabulations were each required to contain at least 100,000 inhabitants. Those intended for agricultural tabulations were each required to contain at least 10,000 farms. Those areas designated as Metropolitan state economic areas were each required to contain at least one central city of not less than 50,000 inhabitants, and the entire Metropolitan area, consisting of the county containing the central city and any contiguous counties satisfying certain population and employment requirements, was required to contain not less than 100,000 inhabitants.

3. The areas were required to follow county lines in all cases, and all State boundaries were also State economic area boundaries.
4. The procedure consisted of the following steps: (a) A tentative delimitation for each State was made, plotting on a county outline map for this purpose regional delineations and type of farming areas that were available from previous work of others; (b) these tentative delimitations were then tested with statistical data in accordance with the above requirements regarding size and homogeneity with respect to specified indices on land use, level of living, population characteristics, type of farm composition, et cetera; (c) the tentative delimitations were thus revised and submitted to other agencies for review—departments of agricultural economics of each State agricultural college, each of the State statisticians of the Crop Reporting Service, and persons or agencies representing a non-agricultural interest.
5. The final review and determination of boundaries were made by a working committee consisting of members from the Census Bureau and from the Bureau of Agricultural Economics.

I have stated what seems to be the primary problem confronting the Bureau—economically meeting demands for enumerated information. Considering this as the main objective, I can readily conceive of the Bureau's staff convincing me that the above rule for delimiting the Areas possesses something that could be regarded as optimal properties. That is to say, the Bureau's technical staff could demonstrate to my satisfaction that the set of areas generated by the above procedure is a "better" set than that obtained, for example, by the different procedure used by Rand-McNally in its partitioning of the nation into a set of areas. "Better" refers to the Bureau's particular purposes or objectives, which in general would not be the same as the objectives that Rand-McNally's areas were designed to serve. There now exists

no systematic method or theoretical basis for analyzing the performance properties of a procedure for obtaining a set of areas comparable, for example, to the theoretical basis that is available for analyzing a sample survey design for obtaining a set of observations upon a population. But the problem is in essence the same. Although there is no known way of applying a formal test of the efficiency with which a given procedure generates a set of areas satisfying predetermined requirements, the Bureau nevertheless would seek a way of determining a set of areas that is optimum in this sense of being economical. The point of view of the Bureau, one may suppose, would be reflected in what Morris H. Hansen—who, incidentally, is chairman of the Bureau of the Census Committee on Statistical Areas—had to say on another occasion and in connection with a different problem. " . . . The over-all test that we [the Bureau of the Census] apply to a sample design is that it shall yield the desired information with the reliability required at minimum cost; or conversely, that at a given cost it shall yield the estimates desired with the maximum reliability."[1] The objectives of the Bureau in the case of these area designations involve presentation as well as estimation; and it is only in such terms of efficient statistical estimation and presentation that I am able to understand the size requirements and the requirement of homogeneity with respect to the prescribed indices as having operational meaning.

II

Up to now I have interpreted the Bureau's efforts as having been directed to the problem of determining an *optimal* set of geographic areas for the particular purposes which the Bureau serves. This is an easily understood conception; but the monograph, in my judgment, confuses the discussion by intermixing, along with such paragraphs as have been quoted above, statements that imply an entirely different idea. At least I get the impression that within the minds of responsible members of the staff advising the Bureau there was an idea corresponding to some kind of "natural area unit" containing something that could be called a "distinct economy" that can be discerned or perceived as an operating "unit." Each State was to have been "divided into its principal units" such that "within each unit a distinctive economy prevails." This notion would seem to imply that "distinctive economy" is a term referring to a natural entity that any competent person can recognize when it is held up before him. In accordance with this idea

[1] Hansen, Hurwitz, and Gurney, "Problems and Methods of the Sample Survey of Business," *Journal of the American Statistical Association*, 41 (1946), p. 174.

of "natural area unit," there would be, not an optimum set of areas for a specified purpose among an indefinitely large number of possible sets of areas, but some certain and existent number of distinct quasi-organic entities; and the procedure that was sought by the Bureau, one may be led to suppose, was a procedure so designed as best to facilitate the discerning and setting apart of these entities. The 501 that were set apart would, correspondingly, not be an estimate of the "best" number but rather an estimate of *the* number. "Because Census data are the sole or principal source of information about many aspects of social and economic life," the monograph states, "there has been an increased demand that the Bureau of the Census undertake to provide more detailed information about the major socio-economic areas in each of the States. The State economic areas have been established in an effort better to serve that growing body of analysts . . . "who are conducting their work in terms of "functionally defined units of area."

The term "functional area" or "functional unit" is used at a number of places to describe the areas, but there is no explanation of what meaning the word "functional" is supposed to convey. When it is stated in the monograph that within each of the area units being sought there prevails a "distinctive economy," an attempt is made to define this latter term. "The term 'economy' is used here in its broadest sense; it refers to the total adjustment which the population of an area has made to a particular combination of natural resources and other environmental factors."

This statement, however, evidently has no operational content. In an empirical field, a definition is supposed to serve the purpose of setting apart those attributes by which that being defined may be recognized and classified under the name assigned to it. The quoted statement offers the following instructions: When confronted with a geographic area containing (1) a population of human beings and (2) a combination of natural resources and other environmental factors, the research worker is to inquire into whether or not (1) has adjusted itself to (2); if adjustment is observed to have taken place, then the geographic area in question is to be classified as containing a "distinct economy." But in accordance with these requirements any area whatsoever inhabited by human beings engaged in exploiting "natural resources and other environmental factors" would qualify as an "area unit containing a distinctive economy," for ordinarily the expression "engaged in exploiting" will do well enough as an equivalent for "adjustment" in this context.

No special meaning is explicitly assigned to "total adjustment," but

one gets the impression that in some way the idea of "homogeneity" is associated with the ideas of "total adjustment" and "functional grouping." The test for "homogeneity" that was improvised by the Bureau involved the following steps:

> The indices [intended as measures of "land use", of the allocation of employment among broad classes of industry, of the "level of living", of population characteristics, of "type of farm composition", of "value of farm products" by type of product, of "crops and yields", of "livestock production", etc.] for all counties of a State were listed on worksheets, grouped according to the tentative delimitation. Sums of indices for each grouping were taken, and the mean value of each index for each group was computed. The index values for each county were compared with the mean values for the two or more groups in which the county could be placed. Wherever there was indication that a county would deviate less from its group mean for any index if it were moved to an adjoining group, the fact was noted. . . . When all indices for a county had been compared . . . the notations of deviation were studied to determine whether or not jointly they indicated that the county should be reclassified. . . Because several index values were being employed to determine homogeneity, the indices were seldom unanimous in placing a county either in one grouping or in another. The indices indicating change and those indicating no change were considered for their relative importance in characterizing the two areas being separated.

This is a plausible procedure for stratifying a population; and it is easily understood how the strata obtained may be found to have useful properties—for example as indicated above, in the improvement of the efficiency with which certain population parameters are estimated or of the effectiveness with which information is classified and presented. But I see no grounds for calling them by names implying more than the term "stratum" implies. I can see no more grounds for calling the strata obtained by this procedure "functional groupings" or "distinctive economies" than if the stratification had been made in accordance with some kind of criterion of "heterogeneity." No explanation is given which would convey an understanding of how the "adjustments" made by the population within any one of these strata may be regarded as different in essence from the "adjustments" made by the population within any other designated area.

This seeking to meet the "needs" of political economists by providing them with "functionally defined units of area" is based upon what I regard as a misunderstanding of the needs of political economists. The need for these area units was described in the monograph as follows (italics added):

> Numerous problems of *maladjustment* to the environment, of *improper* use of resources, of recurring or chronic conditions of unemployment or under-

> employment in certain of these areas ['that have *emerged* as a result of the *adjustment* of the population to the physical environment'] tend now to be regarded as problems of the Nation, and are studied by a variety of specialists. Increased *awareness* of these conditions has accentuated the need for more knowledge about local areas. The growing concern about problems of these types has fostered a general tendency for analysts to descend below the level of State statistics, and to conduct much of their work in terms of smaller, more *functionally defined units* of area.

In my opinion, the important needs of the political economist studying the structure and functioning of an economic system are illustrated by the questions that this statement begs. "Areas" do not "emerge" with this "adjustment" process. The areas have been where they are all along. Something else emerges, and while it is spoken of as an economy or an economic organization, research workers have yet to arrive at a generally satisfying mode of empirically observing and concretely describing its structure and the process by which it is supposed to develop. By what criteria do observers "become aware" of the "improperness" or the "maladjustment" characterizing the performance of an economic system? Certainly not by the mere enumeration of the contents of geographic areas however defined, for these terms imply a comparison and express an evaluation. To perceive "maladjustment" or "improperness" or "underness" or "overness" in an observed situation is to possess the capacity for perceiving and making known by well-defined criteria the qualities characterizing states of "adjustment" and of "properness" in the performance of the organization of economic units.

Work toward establishing this capacity can hardly be said to have begun. There are, to be sure, judgments that are daily made to the effect that per capita incomes in certain areas are *too low* or that industrial development of certain areas is *lagging* or that the structures of the economies of certain areas are *out of balance*. But upon inquiry into the bases of these judgments, one must agree, I believe, that they issue from our passions and sympathies, and one will search in vain for analytical criteria that will stand critical investigation. There are analytical criteria in economic theory by which the consistency of an argument in a social discussion may be judged. But no positive theory now exists which will account for the structural and operating features and developmental processes of a human economy as it performs under acceptable conditions of freedom for individual units and which would assist participants in a social discussion in evaluating how well or how poorly an existing economy is performing and in making explicit the meaning of some such notion as "significant divergence" from "proper performance." The work at this stage is primarily experi-

mental in the efforts that are being made to discover useful ways of describing the structural features and the operating and growth characteristics of an economic system. The Bureau contributes to this work by responding to requests for information. But it cannot inform a research worker in regard to the set of areas he should obtain observations upon any more than it is in a position to specify for administrative agencies what set of areas is "best" for their respective administrative purposes. For a given research worker dealing with some given problem of description or of hypothesis-testing there may exist an optimum set of areas upon which observations are to be made for this piece of research. There exists no unique or single set of "natural area units" as distinct operating entities and component parts of a human economy.

III

The implication of the last statement is that the economy, in its spatial aspects of structure and functioning, is to be regarded as a continuum. It occupies area but it is not to be identified with the area occupied, and it has an objective form of its own as *pattern* and *structure*. The economy is an organization of economizing units. The behavior to be studied is not that of the units but rather of the system of units. As a population system, the organization is characterized by structural and operating features, and these pertain to the system per se and not to any particular set of mortal economizing units that may exist at any specified point in time.

A statistical description of the form and pattern assumed by this system may be made, I believe, without reference to any particular set of geographic sub-areas. For this purpose I would be inclined to develop certain of the ideas of Walter Christaller. Christaller studied what he looked upon as a system of central places.[2] The central places are the cities within an economy, and he classified cities into discrete types, each type having its characteristic function to perform within the system. The pattern that he described was that formed by the spatial orientation of these central places of various types. His principal contribution, I think, lies not in his description of what he thought he saw as a particular pattern but in the steps that he took towards a specification of the attributes of spatial pattern in general. His attributes of pattern include the following: operationally definable types or classes of population clusterings or central places; a natural ordering or gradation of these central places into a hierarchical system of centers;

[2] *Die Zentralen Orte in Suddeutschland*, Jena, 1933.

numerical regularity within this system with respect to the frequency of central places of a given type or size; a numerically specifiable spatial orientation of the different types of central places.

More recently, Bogue, the author of the monograph reviewed above, studied what he regarded as the structure of metropolitan communities.[3] He did not present data for individual metropolitan communities but only average figures for two large classes of communities for each of several parts of the nation. As distance from center is increased, his data describe population density as sloping downward in approximate conformance with a definite rule. Whereas Christaller's conception calls to mind a configuration of population density peaks, Bogue's study suggests a filling in around each peak to form a density configuration that may be conceptually thought of as specifiable in terms of a density function. Thus, the statistical conception of a density function may be employed in this description, and Christaller's notion of a system of central places may be modified into the idea of a spatial density configuration of economic units. This latter would be observed if the locations of the family and individual units at a point in time were to be plotted as coordinates in space, and Christaller's central places would be represented by the points within the area at which localized peak densities are observed.

From this view, the United States, or any other large populated area, appears as an interconnected system of central places. The density of the economic units is at a local peak within each of these central places, but it does not fall off abruptly at anything that can be called an edge or limit of these cities. It declines systematically with distance from the center, and for the major centers the decline continues for a hundred or more miles out, until an area of dominance of an adjacent major center is reached; and then population density rises relatively smoothly to another peak. Passing along connecting links between the major centers, one sees rises to and falls from smaller density peaks at sub-centers. Each major center appears as a hub with spokes, so to speak, extending to other centers. The sub-centers associated with a given major center may be seen as lower order hubs with extending spokes. That which appears as a principal central place from the point of view of a particular set of centers may appear as a subsidiary central place from the point of view of a larger set of places; and the hierarchy of centers scales up finally to a nucleus for the entire system.

[3] *The Structure of the Metropolitan Community—A Study of Dominance and Subdominance*, University of Michigan, 1949.

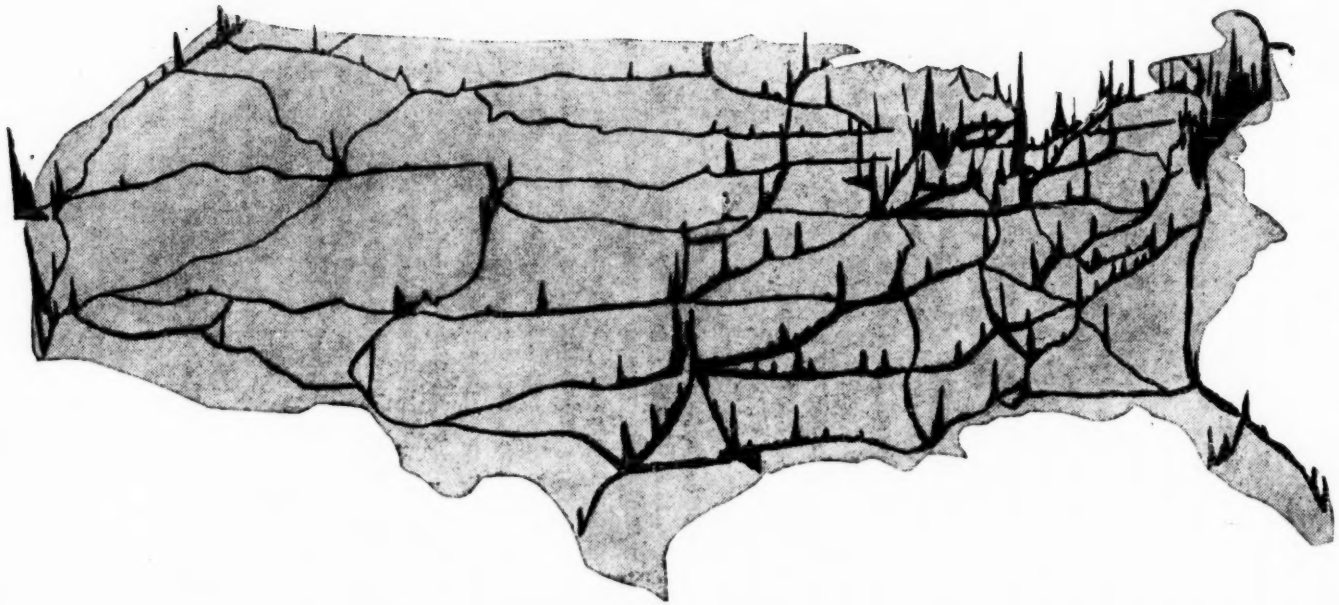


FIGURE I. Relief traffic map. . . . The height of the traffic bands indicates approximately the average density of traffic to be expected at all points on the system.

Source: Interregional Highways, 78th Congress, 2nd Session, House Document No. 379, Government Printing Office, Washington, 1944, between p. 40 and 41.

The connections that link the density peaks into a system consist of the transport and communications networks, and the entire configuration may be visualized without the aid of any sub-area designations regarded as economic regions. Figure I affords a rough visual impression of the configuration covering this nation's area. The densities shown are on a scale such that only the major centers are indicated and refer to highway traffic rather than to the economic units within the nation. The two densities would surely be correlated, however, and this figure illustrates what I have in mind by a spatial density configuration of economic units.

The spacing and clustering of economic units constitutes a spacing and clustering of complementing economic activities. The density peaks represent clusterings of certain types of economic activities, and these central places are formed into systems and sub-systems of central places. This systematic formation may be rationalized and described in common sense terms as follows. Some of the economic units are engaged in activities the locations of which are determined by the locations of natural resources and thus are dispersed over space. The expectation is that each area of substantial size will be endowed with natural resources ranging from those of a type any member of which is found in only a very few places to those of a type the members of which are widely distributed among most areas. All areas contain alike these latter resources, and each contains some one or more of the rarely found resources. The dispersed units are supplied and serviced from central places. In the smallest central places the servicing and supplying operations involve activities the products of which have the smallest market range. The type of value added in these centers is such that there is a relatively large rate of consumption per consuming unit and also such that the rate of production of a tolerably efficient producing unit is relatively small. Among the units serviced and supplied are those whose product specializations may consist of outputs of that type of resources found only in a few places. The mean distance moved from point of origin to destination by the value-added embodied in the products of these units is relatively great. From this extreme, the mean distances covered by resources and products of other types range down to nearly zero; and in these small centers the proportion of units of value-added shipped that terminate at some given distance declines sharply as distance is increased.

Several of these small central places are grouped into a low-order system of places centered upon a somewhat larger place in which are found, in addition to the smallest-range central goods and services,

other central goods and services with a somewhat larger market range. The rate of consumption of these latter central goods per consuming unit will be smaller than that of the goods of the smaller range so that a larger number of consuming units will be required to take a given rate of output. Or the size of the efficient producing unit may be greater than that of those typical of the smaller center so that more consuming units may be required in order to keep a producing unit going than are available within the range of the smallest central place. Also, in this second sized central place the chance is better than for the smallest place that there will be located there some form of processing requiring an assembling of resources not all of which are found within the range of the smallest order of place. Thus, in these second-order centers there are, as in the small centers, economic units engaged in activities producing goods and services with market ranges varying from the relatively large to virtually zero, and for these centers also the proportion of value-added shipped that terminates at some given distance declines sharply as distance is increased. The two types of centers may have similarly shaped distance distributions, and in the short-distance range of the scale the types of products and services are the same in the two distributions. In the longer distance range of the scale the second-order centers will tend to show a somewhat higher degree of product diversification.

Each of the second-order central places has its subsidiary elementary centers about it, and this constitutes the lowest-order system of centers. Several of these are grouped about a third sized place to form a second-order system of centers. This third-order central place will include, in addition to the activities of units providing all the types of central goods and services found in the smaller central places spaced around it, economic activities providing central goods and services with ranges reaching out over the entire system of which it is a nucleus or center. All that was said in comparing the respective distributions of market ranges for the two lower-order types of central places may now be said in regard to this third-order place as compared with the next place below it in the ordering of types of places. A place of this third order is more likely to contain processing activities involving an assembling of resources not found in all areas, and the tendency will be for this assembling to draw from a larger area than is the case for the smaller centers. The products sent the longer distances from these larger centers will typically show a greater diversification than those sent from the smaller centers. Several systems of this higher order are oriented with respect to a still larger central place, and so on; and a

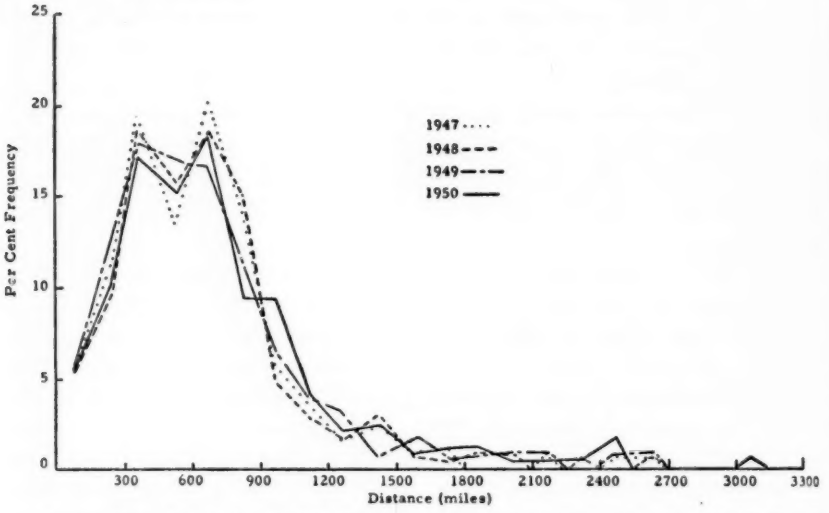
comparison of the situation to be found in any given order of place with that characterizing the next lower order of place would involve the same kind of generalization. A central place of any given order is a source of central goods and services of all the types available in any of the lower-order central places within the system of which it is a center. In addition it is a source of central goods the market ranges of which cover these centers and areas included with this system. And in addition to this it is an assembling and processing point for resources and products peculiar to the territory covered by its system of centers and destined to be shipped to the central places of other systems of centers.

In the above description, the central places are regarded as points of dispersion and points of absorption of economic flows. From each of these centers, value and physical products and services are continually moving outward, counterbalancing the inward flows of value and physical products. For a given center the destination points would form a density configuration and likewise would the origination points for the flows coming into the center. Thus, a second kind of density configuration has been introduced in the description of the structure of the system. The pattern formed by the spatial distribution of the economic units is specified by the first kind. The second kind describes the spatial dispersion of (a) the destinations of the economic flows emanating from any given concentration of units and (b) the originations of the economic flows terminating within any given concentration of units. Suppose that for a given central place and time period the destination were known for each product or value addition produced by the units served by the central place. Then, the distance from the center to the destination for each unit of value-added would be a variate for which a frequency distribution could be formed. This distribution would describe the density configuration of the destinations of the flows originating in the center; and in a similar way there could be constructed a distribution describing the density configuration of the originations of the flows terminating in the center.

As an illustration of what I have in mind as this second kind of density function, I am including here Charts[4] 1–7. These empirical distributions conform only roughly to the idea. They refer to carload rail

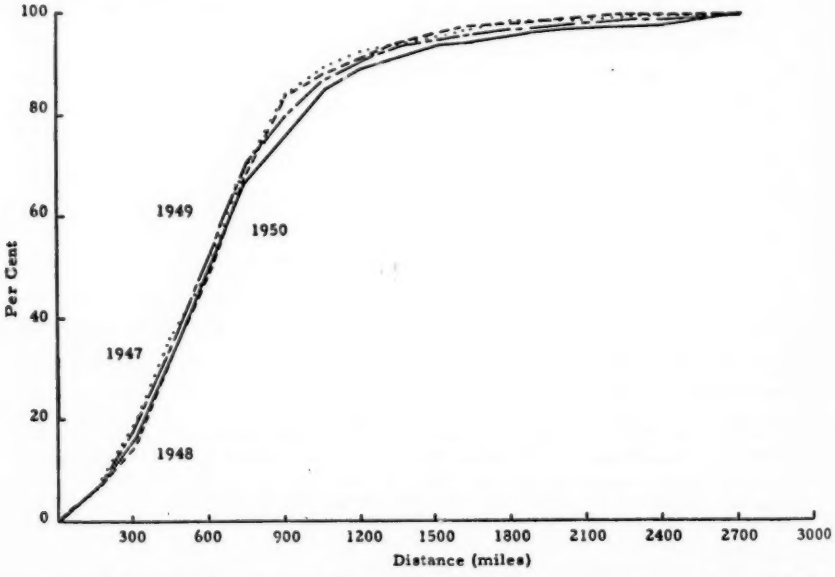
[4] These charts were prepared by Jere Clark with whom I am working in the study of these data and who is preparing a detailed study of economic flow interconnections as these are indicated by the I.C.C. Waybill data. This work is being done under the auspices of the Bureau of Population and Economic Research of the University of Virginia. This Bureau is also making possible a study by Carl Madden pertaining to the development over time of the system of central places represented by the cities of this nation and a study by Carlyle Baskin of the works of Walter Christaller as an illustration of empirical work dealing with the spatial structure of a population system.

CHART 1a. Relative Frequency Polygon Showing the 252 Commodity Groups, by Average Distance Hauled by Rail, United States, First Quarter, 1947, 1948, 1949, 1950.



Source: Data derived from I.C.C. *Carload Waybill Analyses*, Statements No. 517 (1949) and No. 5130 (1950).

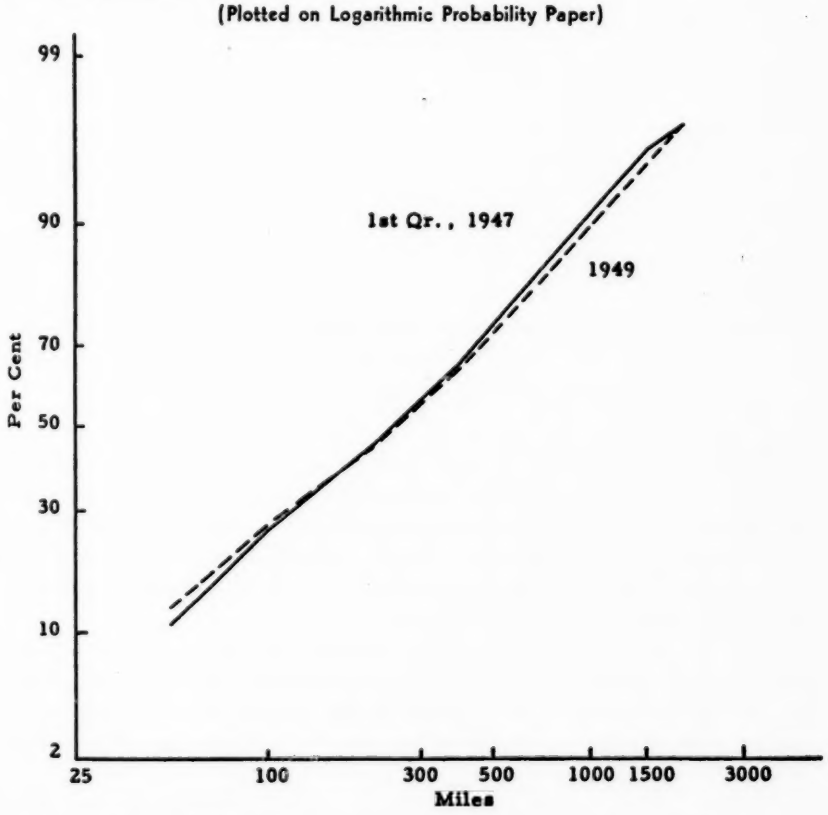
CHART 1b. Cumulative Relative Frequency Distribution of the 252 I.C.C. Commodity Groups by Distance Hauled by Rail, United States, First Quarter 1947, 1948, 1949, 1950.



Source: Data derived from I.C.C. *Carload Waybill Analyses*, Statements No. 517 (1949) and No. 5130 (1950).

tonnage and not to value-added in all products and services, and the originations and terminations are classified by State and not by central place. But they are illustrative nevertheless. Charts 1a and 1b show the distributions for a succession of years of the average distances hauled for some 252 commodity groups. The distributions approximately conform to the same logarithmic normal distribution. It can be shown that from year to year a commodity group will be found within the same general neighborhood of the distance scale, although over longer periods there may be expected systematic shifts of commodity groups along the scale that do not necessarily alter the shape of the entire distribution. This chart represents a classification of commodity groups and indicates a degree of stability in the typing of commodity groups in accordance with average distance hauled; but

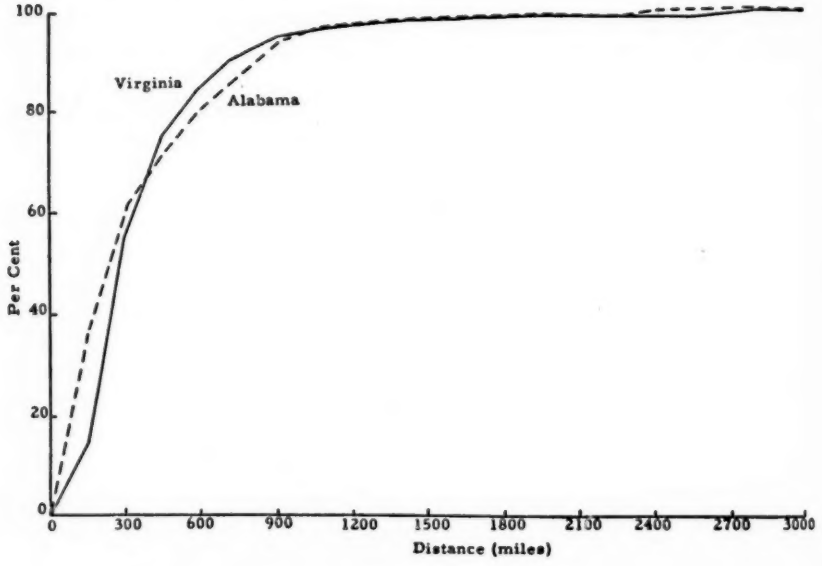
CHART 2. Cumulative Relative Frequency Distribution of Cars of Rail Freight Classified by Distance Hauled, All Commodities, United States, 1949, and 1st Quarter of 1947.



Source: I.C.C. *Carload Waybill Analyses*, Statements No. 4822 (1947) and No. 5111 (1949).

it does not represent the idea of the distance density function referred to in the previous paragraph. Chart 2 presents a different arrangement of the data, and the distributions shown are in rough conformance with this idea of distance density function. It is for the United States as a whole and shows the proportion of total tonnage terminating within any specified distance of the point of origin. The plotting is done in such a way that a straight line indicates a logarithmic normal

CHART 3. Cumulative Relative Frequency Distribution of Carloads of Rail Freight Originated in Alabama and Virginia, by Distance Hauled, Manufactures and Miscellaneous, 1949.



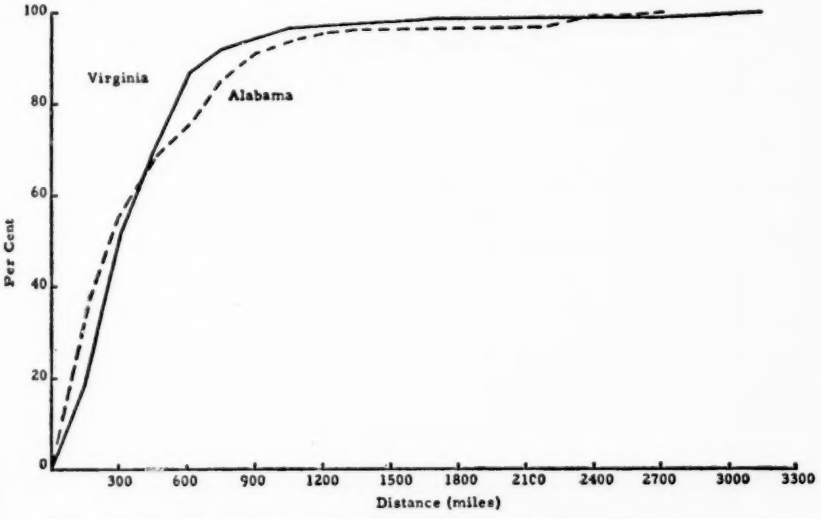
Source: I.C.C. *Carload Waybill Analyses*, Statement No. 5110 (1949).

distribution and it may be seen that the distribution approximately conforms to the same straight line for the periods shown. Two parameters being sufficient for specifying a straight line, this would mean that a knowledge of two parameters is sufficient for the description of the distance distribution of tonnage hauled by rail; and these parameters were stable during the period shown.

When the national distribution is broken into the component State distributions, these State distributions are found to be characterized by essentially the same commodity groups within one end of the scale but by wide divergency in the other end, each State typically having

a commodity specialization in the long distance end. Distributions for Virginia and Alabama are shown in Charts 3, 4, 5, 6, and 7. These charts indicate for given origination points the proportion of traffic terminating within any given distance of those points, the origination points being grouped by States. Various ways of plotting have been employed. Charts 3 and 4 illustrate the similarity between the distributions of the two States. Chart 5, plotted on logarithmic probability paper, shows the similarity between the distribution for a State and

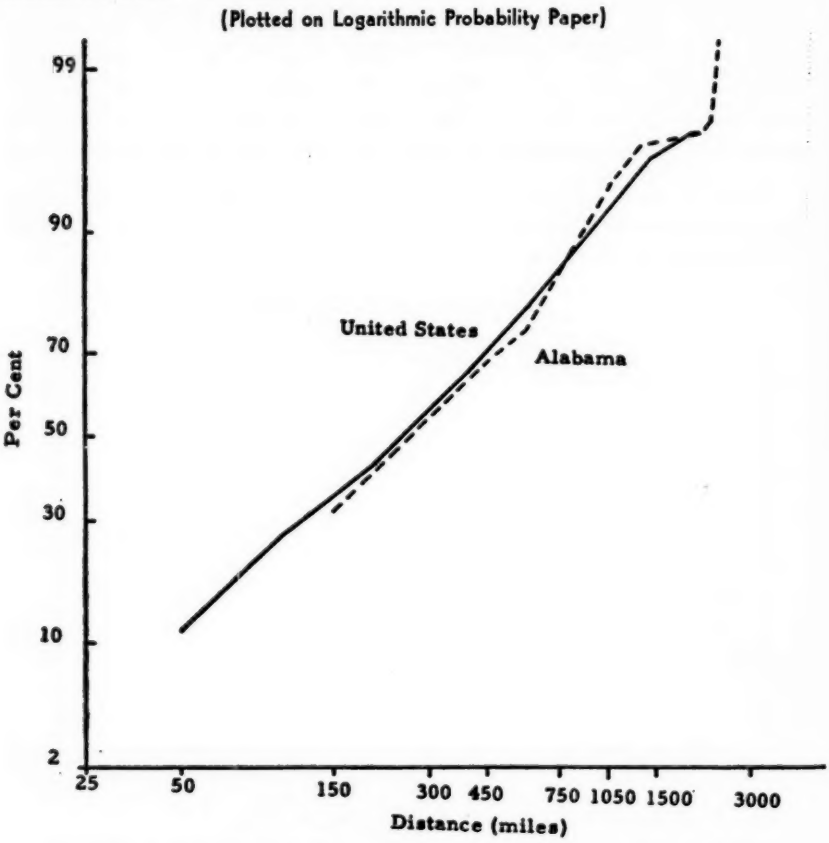
CHART 4. Cumulative Relative Frequency Distribution of Carloads of Rail Freight Originated in Alabama and Virginia, by Distance Hauled, Manufactures and Miscellaneous, 1948.



Source: I.C.C. *Carload Waybill Analyses*, Statement No. 5038 (1948).

that for the Nation. Charts 6 and 7 show logarithmic normal fits to the distributions of a State, one distribution for the terminations of traffic originating in Virginia and one for the originations of traffic terminating in Virginia. Six hundred miles is apparently far enough to include the bulk of the termination points for rail traffic originating at points within a given State and of the origination points for rail traffic terminating at points within a given State. But these distributions refer only to long distance traffic, and a very much shorter distance would be sufficient for including the same proportion of all the value-added, originating or terminating. The relatively stable flow of traffic that can be observed over long distances is stable by virtue of a continual

CHART 5. Cumulative Relative Frequency Distribution of Carloads of Rail Freight Originated in 1948 and in the United States in 3rd Quarter, 1947, by Distance Hauled, All Commodities.

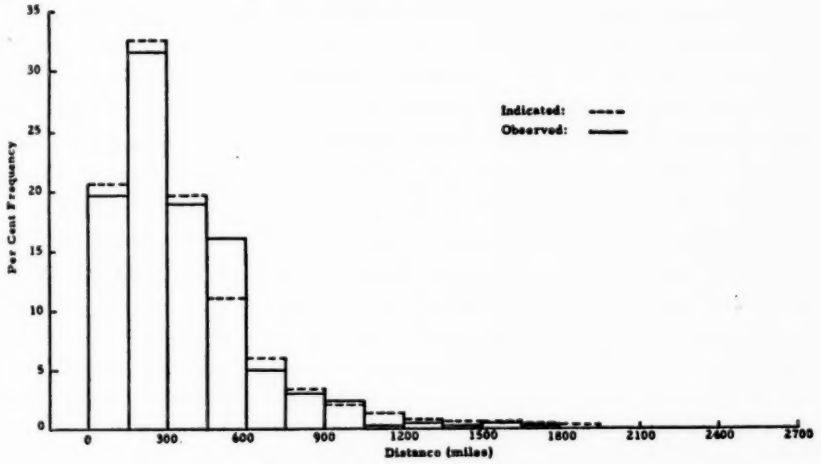


Source: I.C.C. *Carload Waybill Analyses*, Statements No. 4822 (1947) and No. 5038 (1948).

adding-to and dropping-out process, the additions on the average going relatively short distances before dropping out.

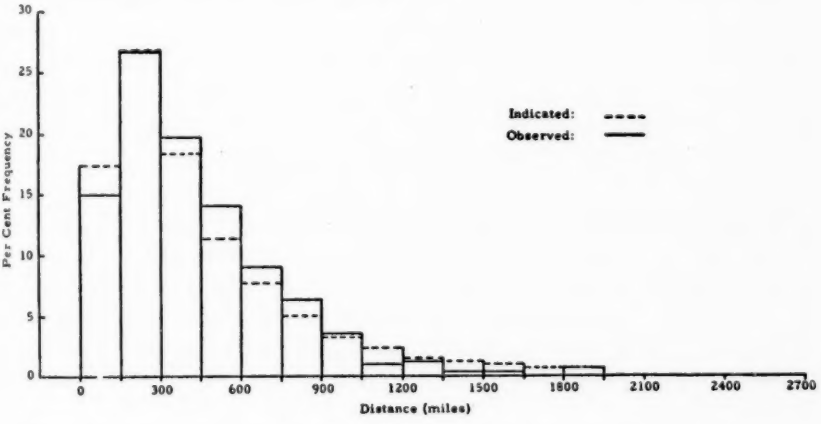
This mode of describing the spatial structure of an economic system does not make essential use of sub-area designations as analytical concepts. When a particular place is under consideration, the essential matter is its orientation, type, and role within the density configuration or system of which it is a part. Its flow connections are shown by its distance distributions—the terminations of its physical and financial flow originations and the originations of its physical and financial

CHART 6. Percentage Frequency Distribution of Observed Number of Cars of Freight *Originating* in *Virginia* Shipped Given Distances, Compared with the Number Indicated by the Logarithmic Normal Distribution, by Distance Intervals, Manufactures and Miscellaneous, 1948.



Source: Data derived from I.C.C. *Carload Waybill Analyses*, Statement No. 5038 (1948).

CHART 7. Percentage Frequency Distribution of Observed Number of Cars of Freight *Terminating* in *Virginia* Shipped Given Distances, Compared with the Number Indicated by the Logarithmic Normal Distribution, by Distance Intervals, Manufactures and Miscellaneous, 1948.



Source: Data derived from I.C.C. *Carload Waybill Analyses*, Statement No. 5038 (1948).

flow terminations. The basic idea in this description is that of a density function; and that part of theoretical statistics which analyzes the conditions that generate distributions having specified forms is relevant in considerations directed to explaining the existence of the particular forms of the distributions observed and the processes by which these forms have developed.

There would seem to me to be little basis for doubt regarding the possibility of constructing a mechanical model that would distribute and from period to period redistribute individual entities among the component small squares of a large partitioned area in such a way as to illustrate certain aspects of the time process by which economic units spread themselves in the course of many decades over an area such as this Nation's. In my opinion the model in conception could be made more realistic than the mechanical models that have been constructed in the study of economic time series phenomena. There are specialists, of course, who can analyze a situation of this sort to the extent that the problem is adequately formulated. The end result would give the expected numbers of elementary areas containing 0, 1, 2, 3, $\cdots$, n individuals after the passage of a time interval of some given length, and the expected distances between density peaks of specified sizes would be indicated. With the continuing passage of time, the system would develop at varying rates in its different parts. The analysis of the process would throw light upon the expected quantitative character of these divergencies in growth rates. The contribution of this would seem to be obvious; for although much of social discussion deals with comparisons and evaluations of regional (or national) growth rates, there is at present an absolute lack of any analytical standard by which to judge critically the observed divergencies in growth rates. "Normal" divergency in growth rates and "normal" concentration of population—concepts implicit in the comparisons made in many instances of social discussion—refer to the "expected" divergency and the "expected" spacing of individuals in a system that is functioning "normally." Analytical knowledge of these latter has yet to be provided; and evaluative comparisons that implicitly assume that in a "normally" working system the different geographic parts of an economic system *should* be developing at approximately the same rates or that there is something "unhealthy" about the functioning of a system in which some parts are showing negative growth rates obviously beg the question of the attributes of a "normally" working system.

LABOR PRODUCTIVITY IN THE SOVIET UNION*

IRVING H. SIEGEL
Twentieth Century Fund, Washington

I

THE productivity of labor, especially in "industry,"[1] is a subject of intense interest in USSR. Its measure is described in an official handbook as the "most important national economic index,"[2] and is projected with production goals in the directives setting forth the five-year plans. The term is probably mentioned more often in Soviet popular and technical literature than in our own.[3] The concept plays a vital role in the "scientific" socialist theory of economic evolution and in the advertised Soviet strategy for achieving the twin objectives of full communism and the end of "capitalist encirclement."[4] Lenin's statement of this role, repeated by Stalin in 1929, has since become a Soviet commonplace:

> In the last analysis, productivity of labor is the most important, the principal thing, for the victory of the new social system. Capitalism created a productivity unknown under serfdom. Capitalism can be utterly vanquished, and will be utterly vanquished, by the fact that socialism creates a new and much higher productivity of labor.[5]

With the adoption of comprehensive "planning" in 1928, a systematic program was inaugurated to supply the material basis for achievement of the long-range Soviet objectives. Agriculture has been collectivized, and many of its operations mechanized. Industry has been rebuilt, expanded, electrified, and otherwise modernized. The urban labor force has been enlarged by streams of "surplus" rural population, and its skill has been progressively elevated by means of on-the-job and voca-

* Revision of papers presented at the December 1951 meeting of the American Statistical Association and at the April 1952 meeting of the New York Chapter. The author is grateful to P. R. Lever, who provided translations of Russian materials in the course of a study conducted in 1949–51 for the Johns Hopkins University Operations Research Office; and to S. Fabricant, W. Galenson, and A. Gerschenkron for their critical comments on the original paper.

[1] Includes principally manufacturing, mining, and electric power supply.

[2] *Slovar Spravochnik po Sotsialno-Ekonomicheskoi Statistike* (Dictionary-Handbook for Socio-economic Statistics), 1948, p. 397.

[3] There is no justification for concluding, however, that productivity is relatively "underemphasized" in U. S. (See S. E. Harris, *Economic Planning*, New York, 1949, pp. 49–54.) The notion of productivity is fundamental to competitive enterprise, but businessmen's decisions concerning it are usually couched in the language of accounting—in terms like "unit labor cost," "unit cost," and "profit." Furthermore, the Soviet press, aptly described by Professor Inkeles of Harvard as a "mass trade journal," reflects by design the state's preoccupation with problems of production and productivity.

[4] On the place of productivity in Soviet theory, see J. H. Towster, *Political Power in the USSR: 1917–1947*, New York, 1948, pp. 410–11, or A. Vyshinsky, *The Law of the Soviet State*, New York, 1948, p. 60. The term "communism" has, since Lenin, been used to describe a "higher" form of "socialism" characterized by distribution of output according to need and the "withering away" of the coercive state.

[5] Quoted by J. Stalin, *Selected Writings*, New York, 1942, p. 135.

tional training. Application of the principle of the "universal obligation to work" has meant the employment of many types commonly found outside the labor force in our own country.

Coercive measures, institutional arrangements, and propaganda direct the energies of the "workers and peasants"—the new "collective masters" of society—toward the prime task of plan fulfillment. The captive trade unions encourage their members to improve skills and increase productivity. Western-style unemployment was "abolished" in 1930. Absenteeism, turnover, and labor mobility are restricted by stern penalties. The regular hours content of the work year has been steadily increased since the early 1930's. The so-called "shortest work day in the world" is officially 8 hours long, and overtime is obligatory. The 6-day work week became standard on the eve of World War II. The relatively small annual output of consumer goods is distributed insofar as practicable on an incentive-pay basis. Honors, prizes, and privileges are bestowed upon exemplary workers. Even before the advent of Stakhanovism in 1935, the techniques of "socialist competition" were stressed for advancing backward workers, for generating "labor enthusiasm," and for hastening maturation of the socially oriented "new Soviet man."

What has been the impact on labor productivity of these and other features of life under Soviet planning—on the course of productivity and the standing of USSR in comparison with advanced "capitalist" countries? To questions such as these the rest of this paper is devoted.

II

It must be recognized at the outset that certain practical and theoretical difficulties preclude the derivation of altogether satisfactory quantitative answers to these questions. Among the practical obstacles are the limited quantity, uneven quality, and inadequate documentation of the Soviet data and indexes relating to production, labor input, and productivity.[6] Virtually full-coverage statistics are preferred, but these are compiled only as an incident to other operations. The personnel involved are generally overburdened, insufficiently trained, and inadequately equipped with mechanical computing instruments.

[6] Much has been written in recent years about the peculiarities of Soviet statistics. See, for example, articles by A. Bergson, C. Clark, M. Dobb, A. Gerschenkron, N. Jasny, S. A. Rice, and others in *Review of Economics and Statistics*, November 1947, February 1948, February 1950, and February 1952; N. Jasny, "International Organizations and Soviet Statistics," *Journal of the American Statistical Association*, March 1950, pp. 48–64; S. A. Rice, "Statistics in the Soviet Union," *Bulletin of the Atomic Scientists*, June 1952, pp. 159–62; M. Mead (and Associates), *Soviet Attitudes Toward Authority*, New York, 1951, pp. 45, 120–21; and the summary of a 1950 "Methodology Conference of Central Statistical Administration USSR," *American Statistician*, June–July 1952, pp. 16–22. O. Morgenstern, *On the Accuracy of Economic Observations*, Princeton, 1950, p. 13, refers to a "report"—most probably apocryphal—that "lie coefficients" were applied by the central statistical authority to data submitted to it during the early 1930's.

The pressure to fulfill plan commitments, moreover, encourages establishments to distort and exaggerate their accomplishments. Authorities decry such abuses, but policing of the reporting system remains lax. The finished statistics are used for administrative purposes and, after proper selection and landscaping, for domestic and foreign propaganda, too.

Of particular interest are the limitations of productivity statistics for the important "industrial" sector and its components, which are called "branches." The available figures, especially those for recent years, refer mainly to output per worker. Since variant, inconsistent, and ambiguous figures have been published for some years, the official industrial productivity series cannot be stated definitively. Since the suppression of detailed data on output, employment, and hours of work began in the mid-1930's, foreign students must generally confine any independent reconstruction of the Soviet record to the pre-war years. Although Soviet analysts made international productivity comparisons in the late 1930's, no similar estimates have been released for the postwar years. In fact, the prewar comparisons are still cited in Soviet literature, and 1929 is still erroneously mentioned as the American peak year.

The formula used for constructing the Soviet index of industrial output per worker has most probably varied, yet the index is commonly supposed to be a quotient of the composite "gross" production and "wage-earner" employment measures.[7] Test computations made by this writer for 1928–35 suggest that some sort of weighted average of the productivity relatives for the major branches was used for at least this interval.[8] A labor-weighted average of branch productivity indexes—the kind of formula sponsored in our own country by technicians of the WPA National Research Project and the U. S. Bureau of Labor Statistics—was officially computed, together with the quotient of the composite production and labor measures, for about five years following May 1943.[9] It is not known which of the alternatives was published for which years.

---

[7] A position somewhat different from the one taken here is maintained by W. Galenson, "Russian Labor Productivity Statistics," *Industrial and Labor Relations Review*, July 1951, pp. 497–98, and "Trends in Soviet Industrial Productivity" (Rand Corporation, P-276), March 6, 1952, pp. 1–4.

The "wage earner" figures refer to *rabochiye*. They may exclude apprentices (*ucheniki*), whose contribution to output is, however, included in the productivity numerator.

[8] See, also, E. L. Granovskii and B. L. Markus, *Ekonomika Sotsialisticheskoi Promyshlennosti* (Economics of Socialist Industry), 1940, pp. 475–79.

[9] *Slovar Spravochnik* comments on the concurrent computation of the labor-weighted average in the 1944 edition, pp. 218–19, but omits reference to it in the 1948 edition, p. 398. This formula was preferred by leading Soviet students but apparently gave disappointing results in the reconversion period. D. V. Savinskii, *Kurs Promyshlennoi Statistiki* (Course in Industrial Statistics), 1949, p. 208, suddenly announced its loss of official status after several pages of favorable comment—as though the information was received near press time.

The gross output (*valovaya produktsiya*) measures which underlie the productivity indexes for industry and its components have been roundly criticized inside and outside USSR.[10] Various features which clearly lead to an unduly favorable picture of Soviet growth have been well publicized—like the inclusion of defective and incomplete goods and the incorporation of new products and new models (especially before 1937) with weights reflecting the price inflation which occurred after 1926–27, the "fixed" base year. But other features which have received less attention also inflate the apparent volume of output and, at least for some years, also exaggerate the changes in production and productivity with respect to the base levels. For example, the output aggregates include the cost of materials and other elements of "gross" price. The extent of duplication depends, not only on the degree of integration and on the reporting practices of establishments, but also on the significance of (net) imports of raw and semi-fabricated materials consumed in industry. Even if the measure of output were neutral to changes in degree of integration, domestic output during the war was doubtless overstated (and the American contribution to it understated) by the inclusion of substantial Lend-Lease shipments of steel and other materials. Similarly, the gains recorded in output and productivity after the war reflected in part the inclusion of materials acquired through trade agreements and as booty, reparations, and shares in the production of joint extraterritorial corporations.

Another source of inflation of output and productivity is the very definition of gross output. The Soviet concept is generally discussed as though it referred only to the "physical" volume of finished goods and semi-manufactures, but it also includes an adjustment for inventory changes and numerous expenditure items presumably reckoned in "fixed" rubles. Among the latter are: communal and other incidental services, internal plant additions, maintenance and repairs, product development cost, contract work, and certain cancelled orders.[11] The opportunities for confusing input with output are obviously abundant in a newly industrializing country which frowns upon markets and establishes prices administratively. Soviet writers themselves have often criticized tendencies to count all work done by "socialist" labor as "productive"; to improve the apparent performance of an establish-

[10] See publications by Gerschenkron and Jasny cited in footnote 6. Among the authoritative outspoken Soviet critics before the war were Sh. Ya Turetskii and A. I. Rotshtein.

[11] See A. I. Rotshtein, *Problemy Promyshlennoi Statistiki SSSR* (Problems of Soviet Industrial Statistics), I (1936), pp. 149–67, or Savinskii, *op. cit.*, pp. 93–100, or *Bukhalterskii Uchet: Sbornik Vazhneishikh Rukvodyashikh Materialov* (Bookkeeping Accounts: Compilation of Most Important Guiding Materials), 1948, pp. 300–06, or *Slovar Spravochnik*, 1948, pp. 112–16.

ment by padding the "other business" account; to record defective or unwanted goods as output; and to include illegally in industrial production any major construction projects executed with plant personnel.[12] Trotsky, who anticipated many of the criticisms made by American students of Soviet statistics, noted particularly the inflation of output aggregates by the inclusion of repair, so that apparent output increases as quality diminishes: "It is not always certain what hides behind [the ruble of output]—the construction of a machine or its premature breakdown."[13]

In 1949, current weights were supposed to replace the largely fictitious 1926–27 prices;[14] and other changes had been made earlier. But the industrial output and productivity indexes remain subject to idiosyncrasies of the gross product concept and will presumably retain much of the buoyancy so desirable from the standpoint of propaganda.

It would be a mistake, however, to suppose that Soviet planners and analysts have been misled. In 1936, industrial production statistics based on another concept (*tovarnaya produktsiya*) were introduced, but not made available publicly. These figures have apparently been used in planning since 1940. The concept is restricted to goods in "finished" or "tradable" form,[15] and the prices are current.

The Soviet gross output definition and measurement practices doubtless lead to industrial production and productivity indexes which rise more spectacularly than conceivable alternative indexes computed by the most preferred Western methods. Conversely, the application of the Soviet concept and methods to American data should lead to indexes rising more steeply than the standard U.S. series. Western practice would favor the "finished" product concept in lieu of "gross"; a chain index for branch output (with or without adjustment of the links for deficiencies of coverage) in lieu of a fixed-base index with

[12] For some recent evidence of dissatisfaction with accounting practices, see A. Arakelian, *Industrial Management in the USSR*, Washington, 1950, especially p. 142; M. Yampolsky, "The Planning of Labour," *Soviet Studies*, July 1952, pp. 78–94; and *New York Times*, August 4, 1952.

[13] L. Trotsky, *The Revolution Betrayed*, New York, 1937, pp. 12–15.

[14] I. A. Sholomovich, *Analiz Khozyaistvennoi Deyatelnosti Promyshlennogo Predpriatiya* (Analysis of Economic Activity of the Industrial Enterprise), 1949, p. 29. According to a recent announcement (*New York Times*, March 13, 1952), production and productivity calculations will be made on the basis of 1952 (in some cases, 1951) prices, and new items will be introduced at comparable 1952 prices. This procedure is consistent with the 1949 change from 1926–27 weights, and hence does not appear to be a new departure.

[15] See, for example, P. Kholodnyi, "Planirovaniye Tovarnoi Produktsii" (Planning of Goods Production), *Planovoye Khozyaistvo* (Planned Economy), 1940, No. 4, pp. 48–52; Savinskii, *op. cit.*, pp. 112–13; and *Bukhalterskii Uchet*, pp. 308–12. Gerschenkron's assumption that the "gross" and "finished" production concepts are close seems untenable in view of the abuses permitted by the former concept and complaints of Soviet technicians leading to the introduction of the latter. See his *A Dollar Index of Soviet Machinery Output, 1927–28 to 1937* (Rand Corporation, R-197), April 6, 1951, pp. 6–7.

pseudo-price weights assigned to new products and models; and, in the event a fixed base is desired, a recent weighting pattern in lieu of a very early one. For industry as a whole, Western practice would favor the combination of branch output indexes with *net* (i.e., value-added or employment) weights, but this preference may not critically affect the results. On the other hand, the characteristic Western treatment of nonhomogeneous output like machinery—either exclusion from the industry total or incorporation by means of a coverage adjustment—could well lead to understatement of the true rates of Soviet industrial output and productivity growth.

Finally, some remarks on the theoretical limitations of the index-number technique are in order. Even if complete, ideal data were available, it would still be impossible to construct unique or universally acceptable temporal indexes for USSR or any other industrializing country undergoing fundamental structural change within a brief period. Index-number statisticians in "mature" Western countries are probably not so concerned as they ought to be over this important fact.

What applies to temporal indexes also applies to "spatial" ones—to international comparisons of multiproduct industry branches or of industry as a whole. Unique or universally acceptable measures cannot be derived for countries which differ significantly in technology, tastes, price patterns, and assortments and specifications of products. Difficult practical problems of valuation also arise when the countries do not engage in substantial trade with each other or with a common third country. In these circumstances, there is a temptation to overstate the significance of comparisons restricted to industry branches with pseudo-homogeneous products, like "steel" or "coal."

III

When the most favorable Soviet claims concerning industrial output per worker are pieced together, a picture of remarkable growth emerges.[16] The 1950 figure was supposed to be 37 per cent above the 1940 level, and the 1951 figure a modest 10 per cent above that for 1950. The 1950 figure was also supposed to be almost 5 times the 1928 index, almost

[16] The series discussed here is based on Sh. Ya Turetskii, *Vnutri-Promyshlennoye Nakopleniye v SSSR* (Intra-Industrial Accumulation in USSR), 1948, p. 30, for 1900, 1940; N. S. Maslova, *Proizvoditelnost Truda v Promyshlennosti SSSR* (Labor Productivity in Soviet Industry), 1949, pp. 27, 42–43, for 1921–39; certain *Pravda* issues for postwar annual link relatives and for the relation of 1950 to 1940; and some other sources for other dates.

The figures for the earlier years refer to "large-scale" industry, to establishments meeting the minimum employment or power criterion set up for the 1913 census. With the development of Soviet industry, the criterion has diminished in importance and appears finally to have been abandoned

7 times that for 1913, and almost 10 times that for 1900. Setbacks are implied for 1940–41 and for the early reconversion period; but none is now shown for any year during the first plan period despite contemporary evidence. The series available for individual branches for the years prior to World War II generally indicate above-average gains for producers' goods and below-average gains for consumers' goods.

Available statistics for output per man-hour show gains of a similar order, except that the rise during the first plan period (1928–32) is sharper and the rise in the third plan period (1938–40) not so steep.[17] If man-hour productivity figures were also reported for the postwar period, they would doubtless show a smaller advance than 37 per cent beyond the 1940 level because the hours content of the work year increased.

Two variant estimates of output per worker derivable for 1950 are somewhat smaller than the figure cited above but are, nevertheless, of the same order. One, based on a lower 1940 level plus the 37 per cent gain for 1940–50, is 4.5 times the 1928 index.[18] The other, based on reported gains of 41 per cent for the first plan period, 82 per cent for the second, and 32 per cent for the third (plus the 37 per cent gain for 1940–50), is 4.6 times the 1928 index.[19]

The average annual increases claimed for Soviet worker productivity are, of course, much greater than the 2 per cent long-term rate commonly cited for our own country. Furthermore, no decisive slackening of the rate is evident through time. During the half century 1900–50, the claimed average annual rate was 4.7 per cent. For the period since the eve of World War I, 1913–50, the average rate was 5.2 per cent. For the period since the inauguration of planning, 1928–50, the average was 7.5 per cent. For the first three plans, 1928–40, the remarkable average of more than 11 per cent is indicated. The rate for the fourth plan, 1946–50, is even higher. Though reattainment of the 1940 level was not claimed until 1948, an index 37 per cent above that level was supposed to have been reached by 1950! The directive for the present

---

[17] Based principally on *SSSR i Kapitalisticheskiye Strany* (USSR and Capitalist Countries), 1939, p. 75, and B. L. Markus, "The Stakhanov Movement and the Increased Productivity of Labour in the USSR," *International Labour Review*, July 1936, p. 7.

[18] The lower 1940 level (3.25 times the 1928 index) is established on the basis of Maslova's 1937 figure and the claim of a 32 per cent increase to 1940. Our higher figure is based on Turetskii's possibly incorrect application of this percentage change to the 1938 figure and his (consistent) assertion that the 1940 index was "more than 3.5 times" the 1928 base (we used 3.6).

[19] These percentage gains for the first three plans, which vary somewhat from other published figures, are still shown in the 1952 article cited in footnote 12.

five-year plan, 1951–55, sets a target of "approximately 50 per cent" for the entire period,[20] or about 9 per cent per year.

In view of the remarks made in the preceding section, these figures overstate the degree of Soviet progress. Starting from a more primitive productivity level, USSR should have exceeded the average American long-term growth rate—especially since the welfare implicit in our high levels of civil liberties, leisure, and consumer goods output does not enter into the measure of production. On the other hand, measurement according to best Western practices would probably have indicated no important gain in Soviet output per worker for the first plan period as a whole; a substantial gain during the second and third periods, say to about 75 per cent above the 1928 level by 1940; and reattainment of the 1940 level by the end of 1950 rather than a significant rise beyond it.

The claims of substantial growth in output per worker during the first and fourth plans—8 per cent per year and about 13, respectively —are unconvincing even though the productivity *potential* was certainly raised during those years. The rise in the potential during the first plan was partly realized in the subsequent years, as the new and renovated plants were utilized more fully and as labor acquired the new "habit of work" and gained experience. In both periods, the influx of workers into the nonagricultural labor force exceeded the planned increase by millions. Such large additions of inexperienced workers would seem incompatible with sharp advances in marginal productivity if output were measured by Western methods. The first plan period, moreover, was very unsettled; it was marked by terror in the agricultural sector, elimination of unemployment compensation, evisceration of trade unions, inefficient management, widespread "wrecking," and acknowledged deterioration of output quality. Output per worker should also have been unfavorably affected by the alleged reduction in official hours of work and in the number of work days per year. Obstacles to true productivity growth in the fourth plan period included poor utilization of the swollen labor force, the reconversion decline in compulsory overtime, and the demoralization incident to the continuation of austerity and the loss of wartime "savings" through currency reform. The official claim that 1950 production exceeded the 1940 performance by 73 per cent doubtless reflected the inclusion of inflationary items of the kind already mentioned— e.g., material imports, the rehabilitation of war-ravaged plants and (illegally recorded) major construction by industry personnel, and the payment of wages to workers retained

[20] *New York Times*, August 23, 1952.

during reconversion for retraining. Measured productivity might also have been favorably affected by asymmetrical treatment of the output and the employment of war prisoners and penalized Soviet citizens engaged in industry.

Among the productivity estimates devised by foreign students for Soviet industry, those of Colin Clark for 1913–36 are probably the best known.[21] These series differ conceptually from the official productivity index; they purport to show net output (in constant American dollars of 1925–34) per employed person (wage earners and others) and per corresponding man-hour. Between the terminal dates, output per employed person supposedly advanced only one-third, while output per man-hour increased by three-fourths. Between 1928 and 1936, the rise in both measures was only about one-fifth. Clark takes no direct account of the new products introduced after 1928; he relies on a small list of products which could reflect the change in the total only by coincidence. Curiously, he estimates a gain of more than two-fifths in both series for 1935–36, the first year of Stakhanovism—a gain about *twice* that claimed by the official Soviet index of output per worker.

Clark has also prepared the longest series available on Russian real national product per man-hour.[22] This series refers to the Western concept of national product rather than to the Soviet concept, which is much less inclusive. It is expressed in American dollars of the period 1925–34 and covers selected years in the interval 1900–47. The figures necessarily show little benefit from the shift of labor from agriculture to industry, since the denominator omits "disguised unemployment" and women in agriculture. Only in 1940 did man-hour productivity exceed the 1913 level—and then, barely. A net gain of only one-eighth is indicated for 1928–40. A serious decline is shown for the first plan period. The 1940 level had not been reattained by 1947.

Jasny's figures for agriculture, which presumably do not exclude "disguised unemployment," suggest a modest rise of 28 per cent in net output per farm worker between 1928 and 1940.[23] The rise, if any, in net output per man-hour was smaller; the increase in the number of work days was probably not offset by the reduction in hours per work

[21] *Review of Economic Progress*, September 1949, p. 1, or *Conditions of Economic Progress*, London 1951, p. 277. See latter source, pp. 185–87, 254–55, for details of the net product index, especially the series which was used in interpolation for early years and extrapolation for 1929–36.

[22] *Review of Economic Progress*, April 1949, p. 2. The treatment of agricultural employment is discussed in *Conditions of Economic Progress*, p. 190; and the productivity series is presented up to 1940, together with various underlying series, on p. 191.

[23] See N. Jasny, *The Socialized Agriculture of the USSR: Plans and Performance*, Stanford, 1949, pp. 420, 676, 714; J. A. Kershaw's review of this book, *American Economic Review*, March 1950, especially pp. 183–84; and Kershaw, "Soviet Agricultural Prospects" (Rand Corporation, P-278), March 7, 1952, pp. 10–12.

day. Since the 1932 level of farm productivity was doubtless much lower than that for 1928, the advance beyond the later year was much more pronounced.

Finally, some increase in worker productivity during the planning era is also indicated for rail transportation.[24] Freight and passenger traffic per "direct" employee in 1932 was 87 per cent above the 1913 average. The increase between 1928 and 1932 must have been of a similar order. A further substantial gain during the second plan period brought the index to 2.7 times the 1913 figure by 1937. The sizable gains during these years reflect the increase in capacity utilization under the same kinds of pressure which elevated American railroad productivity at a still more impressive rate during 1939–43. Soviet rail productivity has been virtually static, however, since 1937, except for the wartime setback. The 1950 goal called for an increase of only 8 per cent above the 1937 or 1940 level, but a rise of only 2 per cent was achieved. In view of the increase in the work week, the 1950 level of rail output per man-hour was doubtless below the 1940 rate.

IV

Despite the defects of the available statistics and the numerous pitfalls of international comparison, there can be no doubt that USSR lags far behind U. S. in productivity. In the final section, it will be suggested further that the gap cannot easily be closed even though the differential might be narrowed.

Soviet researches conducted shortly before World War II concluded that industrial output per worker in 1937, calculated in 1926–27 rubles, amounted to two-fifths the American figure. A slightly higher ratio, 44 per cent, was claimed for output per man-hour. The ratios for output per worker and per man-hour were both supposed to be about one-fourth for 1932; and about one-sixth and one-fifth, respectively, for 1928. USSR had supposedly caught up with Great Britain and Germany in output per worker by 1937, although rates only half as large were claimed for 1928.[25]

These figures might be challenged on various grounds (e.g., the plaus-

---

[24] The railway productivity series was compiled from Soviet sources by J. H. Blackman, of Johns Hopkins University and Rand.

[25] E. Vasilyev and Kh. Kovalson, "Za Dostoinyu Sotsialisticheskogo Obshchestva Proizvoditelnosti Truda" (For Productivity of Labor Worthy of a Socialist Society), *Planovoye Khozyaistvo* (Planned Economy), 1939, No. 3, p. 154; and S. Kheinman, "Ob Ishlishkakh Rabochei Sily i o Proizvoditelnosti Truda" (Concerning Excess Labor Force and Labor Productivity), *Problemy Ekonomiki* (Problems of Economics), 1940, Nos. 11–12, p. 106. The claim of parity with Britain by 1937 is an exaggeration if we accept the British-American ratio of almost one-half computed by L. Rostas, *Comparative Productivity in British and American Industry*, Cambridge (Eng.), 1948, p. 27.

ibility of the dollar-ruble conversion rates). But, even if they are conceded, the gains they show are not remarkable. Comparisons made in 1926–27 rubles should show Soviet accomplishments in a more favorable light than comparisons expressed in dollars or sterling.[26] Furthermore, the years 1932 and 1937 were depressed ones for U. S. and Western Europe. Finally, a different perspective is obtained if a 1908 estimate by Lenin, cited in other contexts by Soviet authors, is juxtaposed to the figures for the planning era. According to Lenin, the productivity of the Russian industrial worker was already 30 per cent of the American in 1908.[27] Since this percentage is higher than the 1928 and 1932 proportions, the alleged gain between 1928 and 1937 largely represents the recovery of lost ground. An advance from 30 per cent in 1908 to only 40 per cent in 1937, a span of 30 years, would seem like slow progress indeed.

Other computations made outside USSR also indicate a large American-Soviet industrial productivity differential. According to Colin Clark, Soviet net industrial output per man-hour (all personnel) was less than one-fourth, and net output per employee was about one-fifth, the corresponding American rates in 1936. These relationships, according to Clark, were less favorable than in 1913. Another set of estimates,[28] used in American government circles during the war, placed Soviet industrial output per man-hour in 1935–38 at 36 per cent of the American rate—a relative performance comparable to the British, slightly inferior to the German, but superior to the Japanese. According to the same source, Soviet productivity in munitions-making was 39 per cent of the American rate in 1944—almost equal to the British percentage, below the German, but substantially above the Japanese.

There also are some postwar estimates, the foundations of which are even less firm. Included among these is a low figure of one-fourth or one-fifth the American industrial output per man-hour,[29] a median figure of 40 per cent the American industrial output per worker,[30] and a broad range of one-fourth to three-fourths the American output per nonagricultural worker.[31]

---

[26] See computations of S. Yugenburg in *Planovoye Khozyaistvo*, 1937, No. 3, pp. 52–54.

[27] Figure cited by B. L. Markus in a volume issued by the Soviet Academy of Sciences, *Proizvoditelnost Truda v Promyshlennosti SSSR* (Labor Productivity in Soviet Industry), 1940, p. 31.

[28] War Production Board figures quoted by R. W. Goldsmith, "The Power of Victory: Munitions Output in World War II," *Military Affairs*, Spring 1946, p. 79.

[29] D. B. Shimkin, "What is Russia's Industrial Strength? II," *Automotive Industries*, August 10, 1950, p. 35, and "Russia's Industrial Expansion," *Fortune*, May 1951, p. 107.

[30] Galenson, P-276, p. 20. Based on an examination of 1937–39 Soviet-American differentials in individual industry branches having output changes measurable in "physical" units; and on probable changes in Soviet and American productivity since then.

[31] J. A. Kershaw, "The Economic War Potential of the USSR," *American Economic Review*, May 1951, p. 479.

In agriculture, too, U. S. enjoys a decisive productivity advantage. Before the war, according to Jasny,[32] over 4 times as many man-days were required per acre of grain in USSR as in U. S., almost 4 times as many per acre of cotton and potatoes, and about 6 times as many per acre of sugar beets. Over 6 times as much labor was expended for the same quantity of milk. Volin notes that the gamut of operations on winter grain required 2.5–3.6 man-days per acre in 1937 in the highly mechanized South Ukraine and North Caucasus, while the average direct American labor requirement for wheat was one man-day or less.[33] This differential in favor of U. S. would be increased if management and other overhead labor were included in the calculations. The persistence of a sizable American advantage in the postwar period is indicated by the sharp advance in American agricultural productivity since 1939 and by continuing Soviet complaints against such irregularities as the employment of superfluous labor and the maldistribution of personnel on collective farms.

When statistics for the whole of productive activity are examined, USSR again appears to be hopelessly outclassed. Clark's series for net national product per man-hour, as we have already noted, omits "disguised unemployment" in agriculture, so it tends to favor USSR. Nevertheless, his figures indicate that Soviet productivity amounted to one-third the American in 1900, one-fifth in 1928, less than one-fifth in 1940, and less than one-eighth in 1947. Both before and after World War II, according to Clark, British and French productivity also greatly exceeded Russian levels.[34] When United Nations figures for 1949 national income in American prices[35] are related to estimates of labor force, we find that Soviet income per labor force participant was only one-sixth the American—about $600 compared to $3500. The Soviet figure was only one-third the British and two-thirds the French.

It may be superfluous to add that Soviet per capita output and consumption standards also lag far behind the American and, in many instances, Western European.[36] USSR clearly has a long way to go to fulfill the "cardinal task," as Stalin has designated it, of outstripping the leading "capitalist" nations in these terms. Soviet literature now

[32] Jasny, *The Socialized Agriculture of the USSR*, p. 443.

[33] L. Volin, *A Survey of Soviet Russian Agriculture*, U. S. Department of Agriculture, August 1951, pp. 33–34.

[34] *Review of Economic Progress*, April 1949, p. 2.

[35] United Nations Statistical Office, *National and Per Capita Incomes of 70 Countries in 1949 Expressed in U. S. Dollars*, New York, October 1950.

[36] Statistics assembled by State Department, United Nations, and Food and Agriculture Organization amply illustrate the Soviet lag. Also of interest is the differential in food purchasing power of an hour's pay, shown by I. B. Kravis, "Work Time Required to Buy Food, 1937–50," *Monthly Labor Review*, February 1951, pp. 147–48. Even the new 1955 goals of Soviet output per capita are only fractions of the corresponding actual 1951 figures for U. S.

speaks of future development as a "gradual transition" from "socialism" to "communism" requiring a vague number of five-year plans.[37]

V

What about the future course of Soviet productivity and the future of the Soviet-American productivity differential? USSR may reasonably be expected to make gains in both respects in the years to come, but rapid progress seems unlikely. There are restraining factors within the very process of broad economic development, even (especially?) under "planning"; in the initial conditions of industrial backwardness and of surplus rural population; and in the incompatibilities of Soviet objectives, ideology, and institutions.

Even under ideal circumstances of abundant entrepreneurial ability at all administrative levels, time would be required for the development of a high productivity potential and of efficient techniques for realization of this potential. Time is needed at best to expand a nation's supply of materials, power, plant and equipment, and adequately skilled labor. It is also needed for the establishment and maintenance of a dynamic, more or less harmonious balance of these expanding factors. In USSR, it will take a long while to achieve the timely replacement of worn-out machinery, the near-optimum utilization of equipment as it becomes available, the specialization of establishments in lieu of excessive integration, the smooth redistribution of industrial (or agricultural) personnel made redundant by mechanization and automatization.[38]

The way in which the surplus labor problem is handled will have an important effect on the rate of Soviet productivity growth. The theory of "universal obligation to work," the fact of rapid shift of labor from agriculture, the absence of unemployment in the market sense, the virtual ban on discharge of personnel, the toleration of labor hoarding, and the frozen status of workers contribute to the accumulation of poorly used reserves in industry as mechanization proceeds. There is no ministry responsible for combing out and redistributing such reserves across ministerial or geographic boundaries—except, in a sense, the police. The existence of these reserves discourages the mechanization of auxiliary activities like loading, unloading, internal transportation, and inspection; it also discourages the most efficient use of mech-

[37] Advancing this "gradual transition" is listed among the duties of communists in the new Soviet Communist Party Statute (*New York Times*, August 21, 1952). In June 1950, an important though inconclusive conference was held on "Means of a Gradual Transition from Socialism to Communism" (translated report in *Current Digest of the Soviet Press*, February 24, 1951, pp. 2–9).

[38] It is significant that the directive of the new five-year plan includes a resolution "to condemn the practice of economic organizations which underestimate the tasks of introducing new technology and mechanization of labor and which permit the incorrect use of manpower" (*New York Times*, August 23, 1952).

anisms installed for these purposes. Of course, the situation may be rationalized in terms of on-the-job training or the registration, control, and maintenance under discipline of a manpower pool for military emergency. If there were such an emergency, USSR could conceivably perform a feat which it failed to accomplish in World War II—the elevation of productivity as the labor force is reduced and as output is maintained or even increased. If "peace" continues, however, disguised unemployment in industry will serve as a brake on productivity advance, and the productivity potential will be far from realized.

Three alternatives seem open to Soviet leaders. One is to redistribute labor compulsorily on a broad scale. The second is to develop a vigorous household sector—recognize leisure as a respectable category, reduce the labor force, and raise the levels of consumption and civil liberties. The third is to allow a surplus to be absorbed gradually in the same plant or in others under the same ministry as the productive economy expands. The third course seems most likely to be followed. The second seems least likely since premature distribution of the advertised ultimate benefits of communism could jeopardize the proletarian dictatorship, weaken control over the population, and hamper development of the Soviet economy along present lines. The first course is impractical, even though it conforms to the popular American conception of the present Soviet way of life.

It is unlikely that "labor enthusiasm"—what Stalin called the "zeal of the millions"—will have a decisive influence on the Soviet productivity level. There is abundant evidence that the Russian worker is not being remolded under heat and pressure into a "Soviet man" who is more productive than an American worker responding to money incentives. In fact, it appears that Stakhanovism, the most significant of the alleged manifestations of labor enthusiasm, largely owes what success it has achieved to incentive wages. Communist leaders still deplore the "chase after the ruble" as a "capitalist remnant" that must be eliminated, but they also extol the "socialist" system of piece rates and even extend it to prison labor.

While USSR struggles to raise productivity, our own should continue to rise at a modest rate. A mechanical extrapolation—based on a growth of 2 per cent per year in American output per worker, 4 per cent for USSR, and a Soviet-American productivity ratio of two-fifths in 1950—would lead to a ratio of only three-fifths by 1970.[39]

Lenin's promise of victory through superproductivity seems impossible of fulfillment. But Lenin also mentioned other routes, and USSR has certainly improved its world position despite inferior productivity.

[39] Galenson, P-276, pp. 24–25, presents six alternative projections, including this one.

ON PROBABILITIES IN BRIDGE

DAN F. WAUGH,*
The National City Bank of New York
AND
FREDERICK V. WAUGH,*
U. S. Bureau of Agricultural Economics

In the game of bridge, the probability of a successful finesse, or a favorable "break" in a suit, depends not only upon the original deal, but also upon all the bids and plays which have been made by one's opponents. The widely quoted *a priori* probabilities, based solely upon the deal, are not generally correct, and are often misleading, after play has begun. This paper shows that the evaluation of probabilities during the course of play of bridge requires the use of the Bayes formula. Two examples are given to illustrate the proper application of that formula to the measurement of probabilities in typical situations arising in bridge.

PROBLEMS of probability arising in the course of play of contract bridge are mentioned in a highly useful and entertaining book [1] written by one of the foremost players in the country. He writes:

> Probability problems in the play of the cards are of two general kinds:
>
> 1. Who holds a particular card?
> 2. How are the cards of a certain suit divided?
>
> The answer to any such question depends on the amount of information you have. Before the play begins, you may make an estimate of the odds; but as you acquire information during the play, your estimate must be revised.

The author does not go on to suggest a method by which the revision could be made.

The writers of the present article follow Borel [2] in suggesting that the Bayes formula be used to evaluate the revised probabilities. Borel uses somewhat artificial examples not taken from actual play. We shall analyze two examples, both taken from actual play.

THE BAYES FORMULA

A general exposition of the Bayes formula may be found in Uspensky [3]. We shall consider its application to bridge, using a notation suited to that purpose. We are concerned with certain probabilities to be determined after the bidding is complete, and play begun. We shall as-

* The authors are indebted to Albert H. Morehead, bridge analyst of the *New York Times*, for many practical suggestions.

sume (as in most bridge columns) that South is declarer. After play has started, South knows at least one, and perhaps others, of the 26 cards originally dealt to his opponents, East and West. He also knows the 13 cards in his own hand, and the 13 cards held by dummy (North). Often he is able to *place* with one opponent or the other, cards other than those which have actually been *played*. If, for example, West chooses to lead a Queen of a suit not bid by East, South may well assume that West also holds the Jack, and probably the 10 also, of that suit. At some stage of play, there are n unplaced cards, r held by West and $n-r$ held by East. These n cards could be partitioned between West and East in any one of $\binom{n}{r}$ different ways. Each of the possible partitions could be defined by naming r cards in the West hand, since if West held these r cards, East would necessarily hold the other $n-r$ unplaced cards. In general, South will need to determine the probability that West's actual holding belongs to a particular group (for example, the group of possible partitions in which West holds the King of spades, or exactly three diamonds).

Let West's possible holdings be classified into the following m exhaustive and mutually exclusive groups:

$$h_1, h_2, \cdots, h_m, \tag{1}$$

and let the capital letters,

$$H_1, H_2, \cdots, H_m, \tag{2}$$

represent the numbers of possible partitions belonging to each of these groups. In general, H_i can be computed by well-known methods of combinatory analysis. The total number of possible partitions is

$$N = \sum_{i=1}^{n} H_i = \binom{n}{r}.$$

If these N possible partitions are considered to be equally likely, the probability that West's holding belongs to h_k is H_k/N. This is, however, a special case. In general, the several possible holdings are not equally likely. In a sense, H_k/N is the *a priori* probability that West's holding belongs to h_k, measured after play has begun but before taking account of the strategy followed by East and West. But H_k/N is quite different from the kind of *a priori* probabilities which are quoted in all bridge books and bridge columns; where the probability of any finesse is always $\frac{1}{2}$, and the probably of a 3-2 break in any 5-card suit is always 0.67826. The *a priori* probabilities that are universally quoted

are based upon an assumed random distribution of 26 cards; 13 to West and 13 to East. The usually-quoted figures are not even accurate measures of *a priori* probabilities after play has begun.

Let S represent the particular series of bids and plays which have been made by East and West, and let

$$(S, H_1), (S, H_2), \cdots, (S, H_m) \tag{3}$$

be the conditional probabilities that S would occur if West's holding belongs to $h_1, h_2, \cdots, h_m$, respectively.

Then, if the hand were played N times, the expected frequency of S from holdings belonging to h_k is $H_k(S, H_k)$; and from all holdings is $\sum_{i=1}^{n} H_i(S, H_i)$. After S has occurred, the inverse probability that West's holding belongs to h_k is

$$(H_k, S) = \frac{H_k(S, H_k)}{\sum_{i=1}^{n} H_i(S, H_i)}. \tag{4}$$

and the inverse probability of each group of holdings is

$$\frac{H_1(S, H_1)}{\sum_{i=1}^{n} H_i(S, H_i)}, \quad \frac{H_2(S, H_2)}{\sum_{i=1}^{n} H_i(S, H_i)}, \cdots, \quad \frac{H_m(S, H_m)}{\sum_{i=1}^{n} H_i(S, H_i)}. \tag{5}$$

In our analysis, we find it convenient to prepare a table, such as

1	2	3	4	5
h_1	H_1	(S, H_1)	$H_1(S, H_1)$	(H_1, S)
h_2	H_2	(S, H_2)	$H_2(S, H_2)$	(H_2, S)
. . .	. . .	. . .	. . .	. . .
h_m	H_m	(S, H_m)	$\dfrac{H_m(S, H_m)}{\sum_{i=1}^{n} H_i(S, H_i)}$	(H_m, S)

(6)

The first column defines each of the m groups of partitions. The frequencies in the second column are computed. The conditional probabilities in the third column must be estimated on the basis of South's observation of the habits of his particular opponents or of the normal strategies of bridge players generally. The fourth column is computed from the data in columns 1 and 2, and the sum shown below. The fifth column exhibits the inverse probabilities defined in (4) and (5), and is

obtained by dividing each entry in the fourth column by $\sum_{i=1}^{n} H_i(S, H_i)$.

Example 1. One of the leading bridge players in this country is Mrs. Helen Sobel. In her book published a few years ago [4], she describes, without giving the bidding, a hand which she played at a contract of six spades. Her opponents' cards were:

spades	Q 9 8 6 4
hearts	Q 10 9 8 7 6 4 2
diamonds	J 8 7 6 5
clubs	A J 9 8 7 6 3 2

West's opening lead was the 4 of hearts. Mrs. Sobel played the Jack from dummy, and it won. She took that as placing the Queen of hearts with West. At trick 2, Mrs. Sobel played the Ace of trump, West and East following suit with the 4 and 6, respectively. East won trick 3 with the Ace of clubs, so that in order to make her contract, Mrs. Sobel had to win all the remaining tricks. East led a small diamond to trick 4. He has no reason to credit his partner with a possible trick in trumps. He can see that the offense has command of hearts and clubs. Diamonds will appear to him the least of several evils. South wins the trick in dummy, and leads the deuce of spades. East plays the 8. South now might play either the Jack, hoping that East holds the Queen, or the King, hoping that of the two spades not yet played, West holds the Queen and East the 9. The "percentage play," according to a well-known writer, was the finesse. Mrs. Sobel, however, played the King; the Queen dropped; and she went on to make her slam.

On what grounds is the finesse described as the "percentage play"? On the grounds that as of the point of completion of the deal, there were $\binom{26}{13} = 10{,}400{,}600$ possible partitions of the 26 East-West cards, 13 to each; and that of these $\binom{25}{13} = 5{,}200{,}300$ were partitions in which East was holder of the Queen of spades; and only $\binom{4}{1}\binom{21}{11} = 1{,}410{,}864$ were partitions in which West was the holder of just 2 spades including the Queen. The $\binom{25}{13}$ corresponds to the fact that all West's 13 cards are selected in this case from the 25 cards *other than* the Queen of spades, which East and West hold between them. The $\binom{4}{1}$ $\binom{21}{11}$

corresponds to the fact that West in this case holds with the Queen, one of the 4 other spades, and his remaining 11 cards are selected from the 21 cards *other than* spades, that the two partners hold. Hence the *a priori* probabilities as usually reckoned were 5,200,300/10,400,600 or 50% that East was holder of the Queen of spades, and only 1,410,864/10,400,600 or 13%, that West held the Queen doubleton.[1]

That was, in Jacoby's language, "an estimate of the odds, before the play begins." It was after East's play to trick 5 that Mrs. Sobel had the decision to make, whether to play the Jack of spades from her own hand, or to play the King. Her preliminary estimate, she is counselled by Jacoby, must be revised in the light of the information that she has acquired in the course of play thus far. At this point, she was able to place the East-West cards as follows:

	West	*East*	*Unplaced*
spades	4	8 6	Q 9
hearts	Q 4	2	10 9 8 7 6
diamonds	8	6	J 7 5
clubs	6	A	J 9 8 7 3 2

We classify West's possible holdings into four groups, according to the spades held, and analyze the probabilities as indicated below:

Set	West's spade holding	H	(S, H)	$H\,(S, H)$	(H, S)
h_1	Q 9 4	3003	1.0	3003	0.467
h_2	Q 4	3432	1.0	3432	0.533
h_3	9 4	3432	0.0	none	0.000
h_4	4	3003	0.0	none	0.000
		$N = 12870$	$\sum H_k(S, H_k) = 6435$		

Mrs. Sobel explains in her book that she had frequently played with the holder of the West hand. It had been her experience that he habitually made an opening lead from worthless trumps in preference to leading away from a suit honor. The opening trick in this hand showed that he had led away from the Queen of hearts in this case. The second trick showed that he held the 4 of the trump suit which he could have led instead. Mrs. Sobel concluded that he would have led away from the heart Queen only as an alternative to leading away from the spade Queen. Those are the only promising cards that East and West have between them, except for the Ace of clubs which East played to trick

[1] We do not say that this is the correct way to compute *a priori* probabilities. Doubtless one should consider the cards already played or placed. But the probabilities shown are those usually used in texts and articles on bridge.

3. Mrs. Sobel in effect rated the East-West plays thus far as substantially certain for any partition belonging to set h_1 or h_2, and as practically certain not to have been made from any belonging to set h_3 or h_4. On those assumptions, the finesse is practically certain to fail, and the play of the king has a better than even chance to succeed. Using equation (4), the probability that West originally held the spade 9 4, or the lone 4 is $[H_3(S, H_3)+H_4(S, H_4)]/\sum_{i=1}^{n} H_i(S, H_i)=0$. The probability that he originally held the Q 4 is $H_2(S, H_2)/\sum_{i=1}^{n} H_i(S, H_i)=0.533$, or 53.3%.

Thus the "estimate of the odds, before the play begins," and the estimate revised according to information acquired during the play, lead to opposite conclusions:

	a priori	*revised*
chance of successful finesse	50.0%	Nil
chance of dropping Queen	13.6%	53.3%

Example 1 illustrates a case in which one of the alternative lines of play is certain to fail. Declarer obviously should try the other alternative, however low the probability of success. Thus, the probability of 53.3% that declarer can succeed in dropping the Queen is of only academic interest. It is shown to indicate that the correct probability differs substantially from the *a priori* probability, which bridge experts use to find the "percentage play." More important is the conclusion that West, rather than East, is certain to hold the Queen of spades. This conclusion can be reached only by an analysis of inverse probability; whether a formal analysis, such as that above, or an informal analysis, such as Mrs. Sobel made.

Example 2. Another leading bridge player, Mr. George Rapee, describes in a newspaper article [5] a match played in the Eastern States contract bridge championships in March 1952.

The bidding was:

South	*West*	*North*	*East*
2 no trump	pass	3 diamonds	pass
3 no trump	pass	5 no trump	pass
6 no trump	pass	pass	pass

After 8 tricks of play, South had been able to place the East-West cards as follows:

	West	*East*	*Unplaced*
spades	Q 7 6 5 2	9 8	
hearts	J 10	3 2	9 5
diamonds	J 9 8 2	K	
clubs	3	5 4 2	Q 9 7 6

The 6 unplaced cards must be distributed 1 to West and 5 to East. Declarer has lost one trick, so must take the remaining five tricks to make his contract. The lead at this point lies in dummy. South can make his contract by leading hearts if West's one unplaced card is the Queen of clubs. Or he can make it by leading clubs if West's unplaced card is one of the two hearts. West's original heart and club holdings may be classified as follows:

Set	Wests' holdings		(S, H)	$H\ (S, H)$	(H, S)
	Hearts	Clubs			
h_1	J 10 9	3	0.2	0.2	0.048
h_2	J 10 5	3	0.0	0.0	0.000
h_3	J 10	Q 3	1.0	1.0	0.238
h_4	J 10	9 3	1.0	1.0	0.238
h_5	J 10	7 3	1.0	1.0	0.238
h_6	J 10	6 3	1.0	1.0	0.238
				4.2	

Success with the club lead depends only on finding the Queen of clubs with East. As of the point of completion of the deal, the probability of that was 0.5.

Success with the heart lead requires (1) that East hold the 9 of hearts, (2) that West hold the Queen of clubs, and (3) that West hold exactly one of the other East-West clubs. As of the point of completion of the deal, the probability of that was approximately 0.008.

If one follows the usual folklore of bridge, relying solely upon *a priori* probabilities, he would have to agree with Mr. Rapee that the heart lead was "the mathematically inferior play" in this case. But this judgment completely disregards the bidding and 8 tricks of play. And the developments during play have greatly changed the probabilities in question.

We now show how the probabilities should be appraised at the beginning of trick 9.

The heart lead will succeed if West's original holding were h_3; the club lead will succeed if West's original holding were either h_1 or h_2. There has been nothing in the East-West play thus far, it appears to us, and we assume that South had the same impression, that has any relevance to the East-West heart-club holdings, *except* West's play of the 10 J of hearts to tricks 2 and 7. South considered it "unlikely," the news item states, that his particular opponent in this case would have "false-carded" by playing 10 and then Jack from a holding of J 10 9. To make it concrete, we assume that West would have done so no more

than 1 time in every 5. We take it as certain that he would not have played 10 J, retaining the 5, from the holding J 10 5.

On these assumptions, as the table shows, the probability that the East-West plays have been made from the holding h_3 is 0.238. The probability that they have been made from either h_1 or h_2 is 0.048. On that reckoning, the heart lead is not "the mathematically inferior play." It has, it is true, less than a $\frac{1}{4}$ chance of succeeding, but the club lead has less than 1/20.

The probabilities of success as ordinarily computed *a priori* as of completion of the deal, and as computed by the method here outlined, as of the point of South's decision, compare as follows:

	a priori	*revised*
Heart lead	0.008	0.238
Club lead	0.500	0.048

CONCLUDING REMARKS

Many additional examples could be given. Books and newspaper columns analyze hundreds of problems of probability arising during the play of bridge. We have never noted a case in which the analysis has been theoretically correct, and find that the answers given are often misleading.[2] Theoretically any problem of probability arising during the play of bridge could be measured correctly by the Bayes formula, and only by the Bayes formula.

The application of the Bayes formula to typical problems arising in bridge may seem too complicated to have practical value in actual play. We note, however: (1) that an understanding of the principles of inverse probability will frequently help a player make a quick and rough judgment as to which alternative is superior, and (2) that one can work out once and for all the probabilities associated with certain typical situations which commonly arise. Concerning the second point, no player would stop the game to compute the *a priori* probability that 6 unplaced clubs would be split 3-3, but he can remember that it is less than $\frac{1}{2}$.

The main difficulty with the analysis is that of making realistic estimates of the conditional probabilities (S, H_k). But a good bridge player can form some shrewd judgments concerning these conditional probabilities by observing the habits of his opponents, and of bridge players generally. Thus East not having a bid against a no trump

[2] We mean no disrespect to the bridge experts. Many of them are brilliant men, and most of them know more about bridge than we do. We object only to the plain fact that the "probabilities" which they universally use do not, in general, measure the probability of success with a given line of play.

contract, West's opening lead very commonly is the fourth highest card in his best suit; against other contracts he may avoid leading away from a King, if he leads a Queen in an unbid suit he probably holds the Q J 10, and so on. Unless one takes account of these habits, conventions, or strategies, he can not estimate which alternative play is superior. The commonly-quoted *a priori* probabilities usually are not very helpful, and often are downright misleading.

REFERENCES

[1] Jacoby, Oswald, *How to Figure the Odds*. New York: Doubleday and Co., inc., 1948, p. 89.

[2] Borel, E., and others, *Traité du Calcul des Probabilités et de ses Applications*. Paris: Gauthiers-Villars, 1940. Part V.

[3] Uspensky, J. V., *Introduction to Mathematical Probability*. New York: McGraw Hill Book Co., 1937. Chapter IV.

[4] Sobel, Helen, *All the Tricks*. New York: Greenberg, 1949. Chapter 2.

[5] Rapee, George, News item in the *New York Times*, March 30, 1952.

PROBABILITIES OF CERTAIN SOLITAIRE CARD GAMES

ROBERT E. GREENWOOD
University of Texas

I. INTRODUCTION

SINCE the days of the Extra Sensory Perception (ESP) experiments, mathematicians and statisticians have been especially interested in "matching" card problems. In this note, two types of solitaire card games involving matches will be considered in which the ordinary 52 card deck is used. For convenience the two card games will be called games A and B. (It will be recalled that the ESP experiments used 5 sequences of 5 cards each, or a 25 card deck. A tabulation of matching probabilities for such a 25 card deck is given by Greville [4]).

In game A with 52 cards, a player plays out the cards one at a time against another deck (or against a standard order) and scores a success if he can go through the deck without getting a match in suit and sequence. In game B the suit feature is disregarded and the deck is considered as a multiple set of four sequences of 13 cards each. The player scores a success if he can go through this multiple deck without getting a match (sequence match only).

Game A is a special case of the classical "problème des rencontres," and has been dealt with by many authors. See Huff [5] and the appended editorial note which gives additional references, or Feller [2], pages 62–63 and 66–69. The probability of obtaining exactly k matches is given approximately by

$$P_k \simeq 1/e \cdot k!, \qquad k = 0, 1, 2, \cdots, 6, 7. \tag{1}$$

Game B has been considered by Battin [1] and Kaplansky [6]. Special generating functions to give the probabilities of the various numbers of matches or hits have been developed. Kaplansky [6, page 911], notes that asymptotic approximations are available for large decks, but states that such approximations are too crude for game B with 52 cards. For a fourfold deck with a large number of cards ($n \gg 52$), the probability of no matches is approximately $1/e^4$.

Joseph A. Greenwood [3] investigated the variance of certain card-matching processes, and from a general formula given by him the variances of the number of matches in games A and B are found to be

$$\sigma_A^2 = 1; \qquad \sigma_B^2 = 64/17. \tag{2}$$

The means of the number of matches are well-known,

$$\mu_A = 1; \qquad \mu_B = 4. \tag{3}$$

II. PROBABILITY OF A WIN FOR GAME B

To calculate the probability P_0 for a win (i.e. no matches) for game B, one may proceed as follows. Let S be defined by the relation

$$S = x_1 + x_2 + \cdots + x_{13}, \tag{4}$$

where x_i is associated with the ith card in some convenient numbering of the 13 cards in a sequence. Then a draw or a play from $(S - x_i)$ could be interpreted as signifying that the ith card was not played but that the other twelve possible cards all had equal chances of being played.

The product

$$\prod_{i=1}^{13} (S - x_i)^4 \tag{5}$$

could be interpreted as signifying that for four specified plays no cards marked 1 were played, for four other specified plays no cards marked 2 were played, etc. Nothing in the above signifies that four cards of each category must be played in the 52 plays. However, if one were select the coefficient of $(x_1 x_2 \cdots x_{13})^4$ in the expansion of form (5), one does get such an assurance. Indeed, this coefficient is the number of ways in which a 52 card deck (considered as a multiple deck of four 13 card sequences) can be played out so as to give no matches against a standard order. The division by the total number of ways of playing out such a deck, $52!/(4!)^{13}$, will give the desired probability P_0. Hence the coefficient of the term $(x_1 x_2 \cdots x_{13})^4$ in the expansion of the form

$$\begin{aligned} &K(x_1, x_2, \cdots, x_{13}) \\ &= \frac{(4!)^{13}}{52!} \prod_{i=1}^{13} (S - x_i)^4 \\ &= \frac{(4!)^{13}}{52!} \left[S^{52} - S^{51}\left(\sum 4x_i\right) + S^{50}\left(\sum 6x_i^2 + \sum_{i \neq j} 16 x_i x_j\right) - \cdots \right] \end{aligned} \tag{6}$$

is the probability P_0. The numerical coefficients in the above expansion are obtainable in terms of multinomial and binomial coefficients.

If one identifies A_i as the absolute value of the coefficient of the term involving S^{52-i} in expansion (6), then

$$P_0 = A_0 - A_1 + A_2 - \cdots + A_{52}. \tag{7}$$

In addition to the probability for no matches (called P_0 above) the probabilities for 1, 2, 3, $\cdots$, matches are of some interest. Rather than develop these explicitly, it is convenient to refer to a procedural method of Kaplansky, who formulated an algebraic symbolism and applied combinatorial methods thereto in order to obtain an analogue to Poincaré's formula in probability. The adaptation of Kaplansky's result [6] to the problem at hand gives

$$P_k = A_k - \binom{k+1}{k} A_{k+1} + \binom{k+2}{k} A_{k+2} - \cdots \pm \binom{52}{k} A_{52}. \tag{8}$$

This is just Poincare's formula with a different interpretation of the symbolism.

The numerical values $\{A_i\}$ may be computed by direct but tiresome methods. One gets

$$A_0 = 1; \qquad A_1 = 4;$$

$$A_2 = \left[\frac{50!}{(4!)^{12}2!}\binom{4}{2}\binom{13}{1} + \frac{50!}{(4!)^{11}(3!)^2}\binom{4}{1}\binom{4}{1}\binom{13}{2}\right]\frac{(4!)^{13}}{52!} = \frac{134}{17}; \tag{9}$$

$$A_3 = 13004/1275.$$

The denominators tend to get larger; to avoid changing denominators, a table of values of CA_i where C was chosen so as to give integral values to CA_i is given in Table 1. A decimal approximation to A_i is also given. The value of C was chosen to be

$$C = 52^{(15)} = 52 \cdot 51 \cdot 50 \cdots 38 = 5{,}860{,}187{,}598{,}276{,}457{,}205{,}760{,}000 \tag{10}$$

(although the least common denominator required was not quite that large).

The estimated values for $A_{16}, \cdots, A_{19}$ were obtained by computing ratios of successive A_i values, $i = 0, 1, \cdots, 15$, constructing a difference table, and extrapolating.

Using truncated forms of relations (7) and (8) one may now estimate the values P_k, $k = 0, 1, 2, \cdots, 7$ as given below in Table 2. Using a rule

TABLE 1

i	CA_i	Decimal approximation to A_i
0	5 8601 8759 8276 4572 0576 0000	1
1	23 4407 5039 3105 8288 2304 0000	4
2	46 1920 6695 1120 3097 3952 0000	7.8823 5294
3	59 7693 1727 6852 5878 4604 1600	10.1992 1569
4	57 1075 3987 0269 2981 1210 2400	9.7450 0200
5	42 9610 1417 7847 3070 1987 8400	7.3309 9640
6	26 4959 3716 6818 1148 6146 5600	4.5213 4624
7	13 7742 2222 3641 6652 5747 2000	2.3504 7462
8	6 1590 4477 5471 4063 5549 6960	1.0509 9789
9	2 4053 2409 1387 4685 5882 7520	.4104 5172
10	8303 3647 3254 8076 8716 8000	.1410 9111
11	2558 1494 8800 1913 1236 3520	.0436 5303 1
12	708 9146 7824 0756 1191 4240	.0120 9713 3
13	177 8606 3073 7964 4982 8864	.0030 3506 72
14	40 6209 3763 9221 2378 4192	.0006 9316 787
15	8 4841 4962 3331 6627 2512	.0001 4477 608
16		.0000 2776 (estimated)
17		.0000 0490 (estimated)
18		.0000 0080 (estimated)
19		.0000 0012 (estimated)

of thumb method, it is possible to get an estimate of the size of the error. Note that $A_i > A_{i+1}$ for $i = 3, 4, \cdots, 14$. The series (8) is an alternating series; arithmetical calculations suggest that for each k in the range above there is an integer m such that after the term involving A_{k+m} the series is an alternating decreasing series. One may then estimate the size of the error as being less than the size of the next (rejected) term. Thus

$$(11) \qquad |\text{Estimated error in } P_k| \leqq \binom{20}{k} A_{20}; \quad k = 0, 1, 2, \cdots, 7.$$

TABLE 2

	Estimated probabilities for game B	Estimated Error	Gram-Charlier probabilities
P_0	.016233	+.00000002	.016161
P_1	.068899	−.0000004	.068953
P_2	.14416	+.000004	.144370
P_3	.19821	−.00003	.198240
P_4	.2013	+.0001	.201113
P_5	.1613	−.0003	.160890
P_6	.1052	+.0008	.105728
P_7	.060	−.0015	.058665

Assuming $A_{20} = 0.0000\ 0002$, the sizes of the estimated errors may be computed from relation (11). These values are also tabulated in Table 2. It must be stressed that these estimates are only useful heuristic guides, i.e. no claim is made that the absolute error satisfies relation (11).

As is well known, game A may be approximately described by the Poisson distribution, relation (1). For game B, however, the mean and the variance of the distribution of hits are not equal (being 4 and 64/17 respectively). Rather than choose the Poisson distribution (for which the mean and the variance are equal) it may be advisable to consider some distribution which does not require equality of these statistics, for example the Gram-Charlier series of type B.[1]

If $f(k)$ is the Poisson distribution entry, then the first two non-zero terms of the Gram-Charlier series [7] are

$$g(k) = f(k) + (1/2)(\text{Variance} - \text{mean})\,\Delta^2 f(k). \tag{12}$$

Values of $g(k)$ have been calculated and are also given in Table 2. The rather close agreement between the Gram-Charlier series and the estimated probabilities from relation (8) is certainly remarkable.

The numerical calculations were either checked by another method or repeated by the same method. The results stated are believed to be accurate, but due to the extreme tediousness of this work no claim for error-free results is made.

III. EXPERIMENTAL TEST

In order to observe the theory in practice, 1000 games of solitaire were played, the number of games being fixed in advance. For each game, the cards were thoroughly shuffled (at least 6 interlacing shuffles and 2 mixed cuts using three or more stacks), and played out for all 52 cards. Two types of scores were kept, one as if game A were being played and the other as if it were game B. For game A, it is possible to compare the distribution of hits with a theoretical distribution based on relations (1), and the variance of the hit distribution with (2). For game B a limited comparison is possible based on the partial distribution given by Table 2, and the variance of the hit distribution as given by (2). The tabulation of results follows:

---

[1] The author is indebted to Nathan Mantel, Biometrics Section, National Institute of Health, for suggesting the Gram-Charlier series and noting its close approximation to the probabilities given by (8).

GAME A

Number of hits	0	1	2	3	4	5	6 and over
Expected games	368	368	184	61	15	3	1
Observed games	359	384	174	66	16	1	0

Expected mean of hit distribution =1, observed mean =0.999. Expected variance of hit distribution =1, observed variance (about expected mean) =0.957.

GAME B

Number of hits	0	1	2	3	4	5	6	7	8	9	10	11 and over
Expected games (from relation 8)	16	69	144	198	201	160	105	60	—	—	—	—
Observed games	14	56	174	194	210	153	108	55	22	8	6	0
Expected games Gram-Charlier	16	69	144	198	201	161	106	59	28	12	4	2

Expected mean of hit distribution =4, observed mean =3.932. Expected variance of hit distribution =64/17 =3.765 approximately, observed variance (about expected mean) =3.466.

It will be noted that the deviations from the expected values are for the most part small and reasonable. The close agreement between the values computed from relation (8) and the Gram-Charlier values has already been commented on.

REFERENCES

[1] Battin, I. L., "On the problem of multiple matchings," *Annals of Mathematical Statistics*, 13 (1942), 294–305.
[2] Feller, William, *Introduction to Probability Theory and Its Applications, Volume 1*. New York: John Wiley and Sons, 1950.
[3] Greenwood, Joseph A., "Variance of a general matching problem," *Annals of Mathematical Statistics*, 9 (1938), 56–59.
[4] Greville, T. N. E., "Exact probabilities for the matching hypothesis," *Journal of Parapsychology*, 2 (1938), 55.
[5] Huff, G. B., "Solution to Problem #4146, derangements," *American Mathematical Monthly*, 53 (1946), 107–10.
[6] Kaplanksy, Irving, "Symbolic solution of certain problems in permutations," *Bulletin of the American Mathematical Society*, 50 (1944), 906–14.
[7] Kendall, M. G., *The Advanced Theory of Statistics, Volume 1, Fourth Edition*. London: Charles Griffin and Co., Ltd., 1948, 154–56.

NON-LINEAR FUNCTIONAL RELATIONSHIP BETWEEN TWO VARIABLES WHEN ONE VARIABLE IS CONTROLLED

R. C. GEARY
Central Statistics Office, Dublin

Joseph Berkson has indicated that when the error of one of two variates is "controlled" in a sense which he defines, and when the true relation between the variates is linear, the classical regression formula can validly be used for the unbiased estimation of the coefficients. In the present paper, Berkson's results are discussed and a theory of estimation and of tests of significance is sketched for the case of controlled experiments when the inherent relation between two variables is non-linear. The results are applied to the study of a constructed example.

JOSEPH BERKSON [1] has made a significant contribution to regression theory and practice in isolating a case in which classical regression theory can validly be used to express a functional relationship between two variables, both subject to error. Berkson's viewpoint justifies the use of linear regression theory without qualification in the typical scientific experiment, in which conditions are partially subject to control. This viewpoint may have been known implicitly or explicitly to many statisticians. The writer was not aware of it, though he has worked on the theory of relationship between statistical variables. That other students may be in the same situation may justify the present further consideration of Berkson's results.

Berkson's main point is that in the case of the "independent variable" it is possible to distinguish between what he terms "controlled" and "uncontrolled" measurements. The objects of the present contribution are (a) to clarify this distinction; (b) to consider the implication of the distinction in its relation to non-linear regression; and (c) in particular, to sketch in a sampling theory, with tests of significance, appropriate to controlled observations. It is shown that, while the R. A. Fisher [2] theories of tests of significance apply unchanged in the linear case, they do not apply in the case of non-linear inherent relationships.

It may be useful, at the outset, to explain briefly the difference between uncontrolled and controlled measurements by reference to a particular example. An operator is making *controlled* observations when he is measuring out quantities x of a drug fixed in advance, say 1 gm, 2 gms, 3 gms, etc., and measuring some reaction v. In making these measurements, however, he makes errors: when he thinks he is measur-

ing 1 gm, he has really measured 0.95 gm, 1.13 gm, etc.; when he thinks he has measured 2 gms, he has really measured 1.81 gm, 2.05 gm, etc. Assuming a linear relationship between the nominal measurement x and the actual reaction v, it is clear that there are two sources of error in v: (i) that due solely to the error in measuring x, and (ii) that in measuring v itself, the actual recorded value being y. The operator's recorded pairs of observations will, of course, be (x, y) where x is 1 gm, 2 gms, 3 gms, etc. In this example, if the inherent relation is linear, i.e. if the relationship between the true values u and v, of x and y respectively, is $v=\alpha+\beta u$, then Berkson has shown that the classical regression formula of y on x will give statistically consistent estimates of the coefficients α and β, making the usual assumption of randomness in regard to the errors of observation in the measurements of the two variables. This conclusion extends to the many-variable case as well, when the relation is linear.

In the case of *uncontrolled* measurements, the operator is not trying to work to any preassigned schedule of measurements of the "independent" variable. He simply measures x and the reaction y and associates one with the other. In this case the classical regression approach gives biased estimates of the coefficient β even if the number of pairs of observations is indefinitely great, for reasons which are well-known: in non-mathematical terms the regression of "y on x" means that one is associating the y's with the x's strung out in order. In ordering the x's, however, one is necessarily taking into account the errors in the x's as well as the "hard core" which is the correct measurement. In simple terms, the x's to the right of the scale will contain large positive errors and those on the left large negative errors. This is why the regression of y on x is "too flat" in the x-direction and, of course, if one orders according to y, too flat in the y-direction in the case of regression of x on y: hence the two regressions. While in Berkson's case of controlled experiment and in the case of uncontrolled experiment, which occurs when neither variable can be regarded as the "independent" variable (case II of Table 1), *both* regressions cannot be valid—in Berkson's case one only, and in case II neither, is valid—Frisch [3] and Reiersøl [7] have shown that in the case of uncontrolled experimentation both regression equations can be used to derive useful information about the true values of coefficients. In the case of two variables in fact, it can easily be shown that the true (functional) coefficient β must lie between the two regression coefficients. Except when the number of pairs of observations is very large and especially when the variables are closely correlated (when the regression lines lie close to-

gether) it is usually as good a procedure as any to take a simple average of the two regression coefficients as the estimate of the true value since this estimate is likely to differ from the true value by less than the error standard deviation from the methods referred to for case II in column 3 of Table 1. While, therefore, the answer to Berkson's rhetorical question [1]: "Are there two regressions?" is "No," it must be added that in many cases valuable information can be derived from the two regression coefficients.

Table 1 shows the appropriate solution applicable to the commonest cases when the inherent relation is linear. The actual observations are always x and y where

$$\begin{aligned} x &= u + e \\ y &= v + f \\ v &= \alpha + \beta u, \end{aligned} \tag{1}$$

e and f are the respective random errors, u and v the respective (unknown) true observations which are presumed to be connected by an exact linear relation. The statistical problem is the estimation of the coefficients α and β.

TABLE 1

INHERENT RELATION LINEAR

Case	Conditions	Solution
I. Classical regression	$e=0$, f statistically independent of $u=x$ and v	Least squares, strictly applicable only when f normally distributed, mean zero and variance the same for all values of x, when the R. A. Fisher [2] significance test theory for coefficients and linearity also applies.
II. Both variables subject to *uncontrolled* error	e and f statistically independent of one another and of u and v	(i) Wald [8]: classification of observations into two groups (ii) Reiersøl [6]: used third variable z related to x and y (iii) Geary [4]: used semi-invariants of two dimensions of total degree higher than 2—applicable only for non-normal distribution in (x, y)
III. Both variables subject to error, that of x controlled	e and f statistically independent of one another and of x and v	Berkson [1]: showed that classical regression formula for y on x applied, subject to special conditions specified above for I. Fisher significance tests apply.

It will be useful for what follows to reproduce what is, in effect, Berkson's algebra. From (1),

$$y = \alpha + \beta(x - e) + f. \tag{2}$$

Let the number of x-grades be m. To simplify the algebra I shall assume that, since the scale of x, the presumed observation, is under the control of the experimenter, he is working to such a scale that

$$\sum_{i=1}^{m} x_i^{2k-1} = 0, \qquad k = 1, 2, \cdots,$$

and further that the "experiment" consists in making m pairs of actual observations (x_i, y_i). There may be many replications of the experiment. Then, denoting by E the average of an indefinitely large number of replications, and from (2), assuming that in each x-grade Ee and Ef are both zero,

$$\begin{aligned} \sum_{i=1}^{m} x_i^{2k} Ey_i = \alpha \sum_i x_i^{2k} = m\alpha\mu_{2k}, \qquad & k = 0, 1, 2, 3, \cdots, \\ \sum_i x_i^{2k-1} Ey_i = \beta \sum_i x_i^{2k} = m\beta\mu_{2k}, \qquad & k = 1, 2, \cdots, \end{aligned} \tag{3}$$

where μ_{2k} is the $2k$th moment of the x, the odd moments being now zero. The simplest solution in (α, β) is

$$\begin{aligned} m\alpha &= \sum_i Ey_i \\ m\beta\mu_2 &= \sum_i x_i Ey_i. \end{aligned} \tag{4}$$

Consistent estimates a and b of α and β respectively are given by

$$\begin{aligned} ma\mu_{2k} &= \sum_{i=1}^{m} x_i^{2k} \bar{y}_i, \qquad k = 0, 1, 2, 3, \cdots, \\ mb\mu_{2k} &= \sum_i x_i^{2k-1} \bar{y}_i, \qquad k = 1, 2, \cdots, \end{aligned} \tag{5}$$

where $\bar{y}_i$ is the mean of the observations y in the ith x-grade; $\bar{y}_i = y_i$ if there are no replications. The classical regression estimates will be obtained fron (5) by taking the lowest values of k. It may be worth observing that, while all values of k will yield consistent estimates of the coefficients, we have no hesitation in favouring the classical regression, or simplest solutions, on grounds of efficency, when the variance of $\bar{y}_i$ is the same in each of the x-grades, for then the regression solutions

are the maximum likelihood solutions (when e and f are normally distributed) and so have minimum variance.

As regards tests of significance and the establishment of confidence limits of the estimates a and b, it is obvious from (2) that the R. A. Fisher [2] theory, with the usual assumptions, applies literally. The error in the ordinate is, however, now $(f-\beta e)$ instead of f as in the case of the Fisher theory. The basic assumptions are satisfied, however, namely that the error variance can be assumed to be the same and the errors normally distributed for the different values of x. Since f and e are statistically independent, the (unknown) error variance will be $\sigma_f^2+\beta^2\sigma_e^2$ and the significance test will be less sensitive and the range of the confidence limits wider than in the Fisher case. In this regard the considerable effect of the coefficient β will be noted.

THE NON-LINEAR CASE

The sampling theory appropriate to the non-linear case of controlled experiments is, however, fundamentally different from the Fisher theory which applies when the error e is zero and, as Fisher has shown, the appropriate tests of significance are basically identical with those used in the linear case. It will suffice to discuss the estimation of coefficients for the case of the polynomial of the third degree. The algebra is extremely simple and the extension to higher degree polynomials will present no special difficulty.

As before, the presumed measure of the abscissa is x but the real measure is u, the error being e, so that $x=u+e$, and the actual measure of the ordinate is y with error f, so that $y=v+f$. The exact relation between u and v is

$$v = \alpha + \beta u + \gamma u^2 + \delta u^3. \tag{6}$$

The problem is to estimate the coefficients α, β, γ, and δ and to discuss their sampling errors. We assume that Ee, Ee^3, and Ef are all zero for each x; that the unknown variance $\epsilon_2=Ee^2$ is the same for all the m values of x; and that

$$\mu_{2k} = \frac{1}{m}\sum_{i=1}^{m} x_i^{2k}$$
$$0 = \sum_i x_i^{2k-1}. \tag{7}$$

In the linear case a sampling theory can be derived from a single observation y for each x. According to the method developed in the

present paper replications are required in the non-linear case.[1] We assume, for simplicity, that the number of observations in each of the mx-grades is the same, i.e. that the whole experiment is repeated n times. From (6)

$$y_i = \alpha + \beta(x_i - e_i) + \gamma(x_i - e_i)^2 + \delta(x_i - e_i)^3 + f_i, \qquad i = 1, 2, \cdots, m. \tag{8}$$

Multiplying (8) across by x_i^0, x_i, x_i^2, and x_i^3 in succession, averaging in each x-grade and summing for all grades, we find

$$\begin{aligned} \nu_{01} &= \xi + \gamma\mu_2 \\ \nu_{11} &= \eta\mu_2 + \delta\mu_4 \\ \nu_{21} &= \xi\mu_2 + \gamma\mu_4 \\ \nu_{31} &= \eta\mu_4 + \delta\mu_6 \end{aligned} \tag{9}$$

where

$$\begin{aligned} \nu_{k1} &= \frac{1}{m}\sum_i x_i^k Ey_i \\ \xi &= \alpha + \gamma\epsilon_2 \\ \eta &= \beta + 3\delta\epsilon_2, \end{aligned} \tag{10}$$

ϵ_2 being the variance of the error e.

From (9) we can readily find γ and δ as follows:

$$\begin{aligned} (\mu_4 - \mu_2^2)\gamma &= \nu_{21} - \mu_2\nu_{01} = \frac{1}{m}\sum_i (x_i^2 - \mu_2)Ey_i \\ (\mu_6\mu_2 - \mu_4^2)\delta &= \mu_2\nu_{31} - \mu_4\nu_{11} = \frac{1}{m}\sum_i (\mu_2 x_i^3 - \mu_4 x_i)Ey_i. \end{aligned} \tag{11}$$

It will be noted that, while ξ and η can also be found from (9), α and β cannot, without the intrusion of the nuisance parameter ϵ_2. Actually the estimates of γ and δ are equivalent to the least squares estimates: those of α and β are not, except in the trivial case of $\epsilon_2 = 0$. We are usually interested only in the formula for δ in order to establish

[1] Theoretically it is possible to conceive of a model whereby, with a *single* y-observation in each grade, maximum likelihood estimates of the error frequency parameters, together with the estimates of the coefficients could be derived. Assume, e.g., that the independent errors e and f are normally distributed with means zero and that the error distributions are the same in each x-grade, and that the number of grades (m) is not less than all the parameters to be estimated. However, even in the simplest non-linear case, i.e. when the coefficient δ in (6) is zero, the five maximum likelihood equations (i.e. those for the estimation of σ_e, σ_f, α, β, γ) are extremely complicated and impracticable of application.

that it is not significantly different from zero. If the estimate of δ, according to the sampling theory developed below, is not significantly different from zero, then the universal variance of y in the ith x-grade will be from (8),

$$(\beta + 2\gamma x_i)^2\epsilon_2 + \gamma^2\epsilon_4 + \phi_2 \tag{12}$$

where ϵ_4 is the fourth moment of the error e and ϕ_2 is the variance of error f. $(m-1)$ estimates of ϵ_2 can accordingly be found by equating the differences of the estimated variances of the sample values of y's in consecutive x-grades to the differences derived from (12), using, of course, the estimates b and c of β and γ, as follows:

$$e_{2i}\{(b + 2cx_{i-1})^2 - (b + 2cx_i)^2\} = \text{var } y_{i-1} - \text{var } y_i, \qquad i = 2, 3, \ldots, m \tag{13}$$

In practice the $(m-1)$ estimates e_{2i} of ϵ_2 will, of course, differ: the simplest procedure may be to take the simple average of the $(m-1)$ estimates as the definitive estimate e_2 of ϵ_2. It is, of course, evident that this procedure for the estimation of ϵ_2 is designed only to furnish a consistent estimate. It is not equivalent to the maximum likelihood solution and is probably not the most efficient. With $\delta = 0$, the estimates a, b, c of α, β and γ are then immediately obtained from (9).

SAMPLING ASPECTS FOR THE NON-LINEAR CASE

It may be remarked at the outset that the approach in this paper to tests of significance in the non-linear case is essentially approximative. It consists simply in making *estimates* of the variances and using these estimates as tests: this test becomes exact only when the number of replications is indefinably large. As is well-known, the Fisher test in the non-linear case is exact, under certain conditions, even when there is only a *single* series of y-observations, i.e. one replication of the experiment.

From (11), the estimates c and d of γ and δ are found by substituting sample means $\bar{y}_i$ for Ey_i and the sampling variances are given by

$$(\mu_4 - \mu_2^2)^2 \text{ var } c = \frac{1}{m^2 n} \sum_i (x_i^2 - \mu_2)^2 \text{ var } y_i \tag{14}$$

$$(\mu_6\mu_2 - \mu_4^2)^2 \text{ var } d = \frac{1}{m^2 n} \sum_i (\mu_2 x_i^3 - \mu_4 x_i)^2 \text{ var } y_i, \tag{15}$$

where n is the number of replications of the whole experiment. If δ is

not significantly different from zero, from (9)

$$\mu_2^2 \text{ var } b = \frac{1}{m^2 n} \sum_i x_i^2 \text{ var } y_i. \tag{16}$$

As to the estimate a of α, its variance, *given* ϵ_2, is, from (9) and (10),

$$\text{var } a = \text{var } \bar{y} - 2(\mu_2 + \epsilon_2) \text{ covar } (\bar{y}, c) + (\mu_2 + \epsilon_2)^2 \text{ var } c, \tag{17}$$

with

$$\text{covar } (\bar{y}, c) = \sum_i (x_i^2 - \mu_2) \text{ var } y_i / m^2 n(\mu_4 - \mu_2^2). \tag{18}$$

STUDY OF A CONSTRUCTED EXAMPLE

To see how the foregoing procedures would work in practice, an artificial example has been constructed on the following lines. Let the inherent relation be

$$v = 3 + 2u - u^2 \tag{19}$$

and let there be 5 $(=m)$ values of x, namely -6, -3, 0, $+3$, $+6$. There were 10 $(=n)$ replications of the whole experiment. The observed values of x, deemed to be these five values, are subject to errors which are assumed to be unbiased. The ordinates are also subject to random error. Actually the "induced" errors were normally distributed with mean zero and variance unity and were in fact those $(\times 1/10)$ given by E. S. Pearson [5, p. 349] for the normal population, the 50 values in the first column of Pearson's table being regarded as the errors e in x and the 50 values in the second column being those (f) for y. Thus the actual values in the case of the second observation (i.e. $x = -3$) of the first of ten experiments were *not* the true values $x = -3$ and $v = 3 + 2 \times -3 - (-3)^2 = -12$ but $u = -3 + 1.0 = -2.0$ and $y = 3 + 2 \times -2.0 - (2.0)^2 - 0.4 = -5.40$. The 50 actual values of y as "observed" are given in the table, with computed means and variances. The theoretical values, which would, of course, be unknown in actual experimentation, are also given as well as two theoretical series of values of the means, one biased due to the error term $-(e)^2$ the mean of which is unity, and the other with this term eliminated. The sample average and variance for $x = -6$ goes far astray because of the abnormally high entry at experiment X due in turn to the fact that in this case the x-error was 3.0. One would probably decide to eliminate or to adjust this value by reference to the other nine values in the column: though this is always a hazardous procedure if bias is to be avoided. If this discrepant value were eliminated, the resulting (sample)

mean would be 44.22 and the variance 177.65. For the following computations the original values have been retained.

The first fact to note is the great variability of the variance from grade to grade, which would make the assumption of equal variance, necessary for the application of the Fisher sampling theory, quite unreal. The value of d, the estimate of δ, the coefficient of u^3, is, using (11), 0.0127. As the standard deviation, estimated from (15), is 0.0292, we infer that δ is not significantly different from zero. The estimates of γ and $\beta\pm$ their standard deviations estimated from (14) and (16) are

$$\begin{aligned} c &= -1.03 \pm 0.12 \\ b &= 2.44 \pm 0.49 \end{aligned} \tag{20}$$

so that the sample values do not differ significantly from the true values -1 and $+2$ respectively.

TABLE 2

CONSTRUCTED EXAMPLE: INHERENT RELATION QUADRATIC

Experiment	Observed values of y for $x =$				
	−6	−3	0	+3	+6
I	−52.25	−5.40	4.86	−4.24	−28.14
II	−57.91	−34.44	1.76	3.36	−14.24
III	−39.15	−15.56	5.51	1.10	−9.36
IV	−42.59	−8.55	4.16	−5.74	−10.09
V	−14.79	−5.59	3.50	1.96	−21.20
VI	−35.89	−9.04	4.01	1.26	−20.70
VII	−40.16	−4.34	−2.80	1.11	−8.26
VIII	−59.40	−17.19	2.70	−0.94	−33.59
IX	−55.84	−17.79	3.36	4.29	−29.29
X	−95.50	−14.74	5.40	−4.79	−17.39
Mean					
Experiments (biased)	−49.35	−13.27	3.25	−0.26	−19.23
Theoretical*					
Biased	−46	−13	2	−1	−22
Unbiased	−45	−12	3	0	−21
Variance					
Experiments (9 d.f.)	440.49	81.25	5.88	12.40	80.27
Theoretical*	199	67	7	19	103

* Unknown in the case of actual experimentation.

As to the estimate of α, it is first necessary to estimate the variance ϵ_2 of the error of observation e of x. Using (13), with the sample values for b and c, the four estimates of ϵ_2 (i.e. e_{22}, e_{23}, e_{24}, and e_{25}) are respectively 2.48, 1.10, 0.81, and 0.80, the mean of which is 1.3, which is taken as the estimate of ϵ_2: the actual value, as we know, is 1. Using (9) and (10)

$$\begin{aligned} &\text{Mean } y = a + c(\epsilon_2 + \mu_2) \\ &\text{or } -15.77 = a - 1.03(1.3 + 18) \end{aligned} \tag{21}$$

giving $a = 4.11$, the actual value being 3. The variance estimate, computed from (17), using sample values, is 1.91, so that the standard deviation is 1.4. The error of estimate of α, namely 1.11, is accordingly less than the estimated standard deviation.

REFERENCES

[1] Berkson, Joseph, "Are there two regressions?" *Journal of the American Statistical Association*, 45 (1950), 164–80.

[2] Fisher, R. A., "Applications of Student's distribution," *Metron*, 5 (1925), 90–104.

[3] Frisch, R., *Statistical Confluence Analysis by Means of Complete Regression Systems*. Oslo: Universitetets Økonomiske Institut, 1934.

[4] Geary, R. C., "Inherent relations between random variables," *Proceedings of the Royal Irish Academy*, 47, A.6. (1942), 195.

[5] Pearson, E. S., "A comparison of β_2 and Mr. Geary's W_n criteria," *Biometrika*, 27 (1935), 333–52.

[6] Reiersøl, O., "Confluence analysis by means of lag moments and other methods of confluence analysis," *Econometrica*, 9 (1941), 1–16.

[7] Reiersøl, O., *Confluence Analysis by Means of Instrumental Sets of Variables*. Uppsala, 1945.

[8] Wald, A., "The fitting of straight lines if both variables are subject to error," *Annals of Mathematical Statistics*, 11 (1940), 284–300.

COMBINATION OF NEIGHBORING CELLS IN CONTINGENCY TABLES

C. C. Craig
University of Michigan

In the application of the χ^2-test for independence to contingency tables in case the expected frequencies in some of the cases are small, a commonly recommended procedure is to coalesce two or more rows and/or columns. Based on Cramér's modified χ^2 minimum method, this paper develops means of combining individual cells instead of whole rows or columns. Specific results are given for a pair of cells in the same row (or column), a block of cells, two pairs which form a 2×2 block, and two row pairs whose columns do not overlap. Results generally are simple if the sets of cells combined form a block or blocks which do not overlap. The complications that arise in the case of partial overlapping are illustrated for two row pairs which have a single column in common. The estimates of the probability of an item falling in each of the three columns on the null hypothesis are obtained from a quadratic equation which, however, always leads to a unique set of consistent estimates.

One of the perennial difficulties encountered by persons who try to follow the directions given in the numerous books on statistical methods in their own research is how to proceed with contingency tables larger than 2×2 in which the expected frequencies in some of the cells are below the deadline of 5 or 10 (depending on the book). Frequently a worker is very reluctant to make the frequently recommended sacrifice of coalescing two or more whole rows or columns. It is beside the purpose of this note to summarize or even comment on the various investigations and discussions of the problem of small cell frequencies. But it may be useful to mention that the exact test available for the 2×2 table in such cases can be extended to larger tables if one is willing to carry out the necessary computations, as has recently been shown by Freeman and Halton [1]. Also recent papers by Adler [2] and Cochran [3] will be helpful in understanding the present status of knowledge and opinion on this issue.

Not long ago my attention was called to a young woman who resolved the difficulty for herself by the very simple procedure of combining a single cell with a too small expected frequency with an adjacent cell, treating the combination as a single cell and then computing a χ^2 in the obvious mechanical way, reducing the degrees of freedom by one. It is no longer clear that the χ^2 obtained in this way has (asymp-

totically) a χ^2-distribution with the degrees of freedom found in this way. This set me to thinking about the possibility of a correct procedure of this kind developed from an acceptable mathematical model.

The results which follow were derived by what Cramér has called the modified χ^2 minimum method. For full details and in particular for a proof that a χ^2 obtained in this way has asymptotically a χ^2 distribution, the reader is referred to Cramér's book [4]. Briefly, for an $r \times s$ contingency table, under the null hypothesis of no association or of independence, the probability p_{ij} that a randomly chosen individual belongs simultaneously in the ith row and jth column of the table ($i=1, 2, \cdots, r$; $j=1, 2, \cdots, s$) is the product of the probabilities $p_{i.}$ and $p_{.j}$ that an individual falls into the ith row and the jth column respectively. Since

$$\sum_{i=1}^{r} p_{i.} = \sum_{j=1}^{s} p_{.j} = 1,$$

there are $r+s-2$ unknown parameters to be estimated from the data before the expected cell frequencies can be found from which a χ^2 can be computed. As Cramér remarks the modified χ^2 minimum method is equivalent to the use of the principle of maximum likelihood for the estimation of the $p_{i.}$ and $p_{.j}$. If the observed frequency in the cell at the intersection of the ith row and the jth column is ν_{ij}, then the probability of occurrence of the observed cell frequencies on the null hypothesis is

$$C(p_{1.}p_{.1})^{\nu_{11}}(p_{1.}p_{.2})^{\nu_{12}} \cdots (p_{r.}p_{.s})^{\nu_{rs}} = Cp_{1.}^{\nu_{1.}} \cdots p_{r.}^{\nu_{r.}} p_{.1}^{\nu_{.1}} \cdots p_{.s}^{\nu_{.s}} \tag{1}$$

in which

$$\nu_{i.} = \sum_{j=1}^{s} \nu_{ij}, \qquad \nu_{.j} = \sum_{i=1}^{r} \nu_{ij},$$

$$\sum_{i=1}^{r} \nu_{i.} = \sum_{j=1}^{s} \nu_{.j} = n,$$

and C depends on the ν_{ij}'s but not on the $p_{i.}$'s nor on the $p_{.j}$'s. In this case the method of maximum likelihood consists of determining the values $\hat{p}_{i.}$ and $\hat{p}_{.j}$ of $p_{i.}$ and $p_{.j}$ respectively for which the expression (1) is maximized for fixed values of the marginal totals, $\nu_{i.}$ and $\nu_{.j}$, and then using those values as the sought estimates of the $p_{i.}$ and $p_{.j}$. The equations which have the $\hat{p}_{i.}$ and $\hat{p}_{.j}$ as their solutions are

$$(2)\qquad \begin{aligned} \sum_{j=1}^{s}\left(\frac{\nu_{ij}}{p_{i\cdot}}-\frac{\nu_{rj}}{p_{r\cdot}}\right)&=\frac{\nu_{i\cdot}}{p_{i\cdot}}-\frac{\nu_{r\cdot}}{p_{r\cdot}}=0, \qquad (i=1,\cdots,r-1)\\ \sum_{i=1}^{r}\left(\frac{\nu_{ij}}{p_{\cdot j}}-\frac{\nu_{is}}{p_{\cdot s}}\right)&=\frac{\nu_{\cdot j}}{p_{\cdot j}}-\frac{\nu_{\cdot s}}{p_{\cdot s}}=0, \qquad (j=1,\cdots,s-1)\end{aligned}$$

which give the usual estimates

$$\hat{p}_{i\cdot}=\frac{\nu_{i\cdot}}{n}, \qquad \hat{p}_{\cdot j}=\frac{\nu_{\cdot j}}{n}.$$

To apply this method to the case in which a pair of neighboring cells have been combined, we assume that $s>2$ and that the cells combined are the first two in the first row. That is, $\nu_{11}+\nu_{12}=w$ is known but not ν_{11} and ν_{12} separately. The null hypothesis to be tested becomes: $E(\nu_{ij})=np_{i\cdot}p_{\cdot j}$ for $i=1,\cdots,r, j=3,\cdots,s$, and $E(w)=np_{1\cdot}(p_{\cdot 1}+p_{\cdot 2})$. Now the modified χ^2 minimum or maximum likelihood equations corresponding to (2) become for rows

$$\frac{\nu_{i\cdot}}{p_{i\cdot}}-\frac{\nu_{r\cdot}}{p_{r\cdot}}=0, \qquad (i=1,\cdots,r-1)$$

as before, while for columns they are

$$(3)\qquad \begin{aligned} \frac{w}{p_{\cdot 1}+p_{\cdot 2}}+\frac{\sum_{i=2}^{r}\nu_{i1}}{p_{\cdot 1}}-\frac{\nu_{\cdot s}}{p_{\cdot s}}&=0,\\ \frac{w}{p_{\cdot 1}+p_{\cdot 2}}+\frac{\sum_{i=2}^{r}\nu_{i2}}{p_{\cdot 2}}-\frac{\nu_{\cdot s}}{p_{\cdot s}}&=0,\\ \frac{\nu_{\cdot j}}{p_{\cdot j}}-\frac{\nu_{\cdot s}}{p_{\cdot s}}&=0, \qquad (j=3,\cdots,s-1).\end{aligned}$$

The row system solves as before to give

$$\hat{p}_{i\cdot}=\frac{\nu_{i\cdot}}{n}, \qquad (i=1,\cdots,r).$$

To solve the column system we first multiply the jth equation by $p_{\cdot j}, j=1,\cdots,s$, and then add, getting

$$n=\frac{\nu_{\cdot s}}{p_{\cdot s}},$$

so that

$$\hat{p}_{\cdot j} = \frac{\nu_{\cdot j}}{n}, \qquad (j = 3, \cdots, s).$$

Finally, the first two equations of the column system are solved separately, giving

$$(4) \qquad \hat{p}_{\cdot 1} = \frac{(\nu_{\cdot 1} + \nu_{\cdot 2})(\nu_{\cdot 1} - \nu_{11})}{n(\nu_{\cdot 1} + \nu_{\cdot 2} - w)} \quad \text{and}$$

$$\hat{p}_{\cdot 2} = \frac{(\nu_{\cdot 1} + \nu_{\cdot 2})(\nu_{\cdot 2} - \nu_{12})}{n(\nu_{\cdot 1} + \nu_{\cdot 2} - w)}.$$

With these estimates of the $p_{i\cdot}$ and the $p_{\cdot j}$, the expected cell frequencies and χ^2 are calculated in the usual way. The resulting χ^2 has $(r-1)(s-1)-1$ degrees of freedom. The reader may have noted that the expression for $\hat{p}_{\cdot 1}$ is the product of the proportion of the total frequency falling into the first two columns by the proportion of items in the first column to the total in the first two columns after the two cells that were combined have been deleted. It is obvious that the method here developed for the first two cells in the first row applies equally well to any two cells lying in the same row or in the same column; if it seems simpler, one may first interchange rows and columns in the table until the cells to be coalesced lie in the position for which the formulas have been given.

As an example, consider the following 4×3 table

				Totals
18	29	70	115	232
17	28	30	41	116
21		11	20	52
113		111	176	400

constructed by combining the first two cells in the third row in an example used in Hoel's textbook [5] for which both expected frequencies are less than 10 on the null hypothesis of independence. We have

$$\hat{p}_{1\cdot} = \frac{232}{400}, \qquad \hat{p}_{2\cdot} = \frac{116}{400}, \qquad \hat{p}_{3\cdot} = \frac{52}{400},$$

and

$$\hat{p}_{\cdot 3} = \frac{111}{400}, \qquad \hat{p}_{\cdot 4} = \frac{176}{400},$$

as usual. But now,

$$\hat{p}_{\cdot 1} = \frac{113}{400} \cdot \frac{35}{113 - 21} = 0.1075^{-}, \qquad \hat{p}_{\cdot 2} = \frac{113}{400} \cdot \frac{57}{113 - 21} = 0.1750^{+}.$$

Then

$$E(\nu_{11}) = 400\hat{p}_{1\cdot}\hat{p}_{\cdot 1} = 24.9,$$

etc., for the single cells while

$$E(w) = 400\hat{p}_{3\cdot}(\hat{p}_{\cdot 1} + \hat{p}_{\cdot 2}) = 14.7.$$

Then from the 11 cells, $\chi^2 = 17.9$ which for $3.2 - 1 = 5$ degrees of freedom lies between the 1 per cent and 0.1 per cent points.

This method is readily extended to the case in which the coalesced cells form a rectangular block. For example, suppose that only the total $w = \nu_{11} + \nu_{12} + \nu_{21} + \nu_{22}$ of the 2×2 square in the upper left corner of the $r \times s$ table ($r \geqq 3$, $s \geqq 3$) is given. One now has maximum likelihood equations of the form (2) for both rows and columns but the solution proceeds as before, giving,

$$\begin{aligned}
\hat{p}_{1\cdot} &= \frac{(\nu_{1\cdot} + \nu_{2\cdot})(\nu_{1\cdot} - \nu_{11} - \nu_{12})}{n(\nu_{1\cdot} + \nu_{2\cdot} - w)}, \\
\hat{p}_{2\cdot} &= \frac{(\nu_{1\cdot} + \nu_{2\cdot})(\nu_{2\cdot} - \nu_{21} - \nu_{22})}{n(\nu_{1\cdot} + \nu_{2\cdot} - w)}, \\
\hat{p}_{\cdot 1} &= \frac{(\nu_{\cdot 1} + \nu_{\cdot 2})(\nu_{\cdot 1} - \nu_{11} - \nu_{21})}{n(\nu_{\cdot 1} + \nu_{\cdot 2} - w)}, \\
\hat{p}_{\cdot 2} &= \frac{(\nu_{\cdot 1} + \nu_{\cdot 2})(\nu_{\cdot 2} - \nu_{12} - \nu_{22})}{n(\nu_{\cdot 1} + \nu_{\cdot 2} - w)}, \\
\hat{p}_{i\cdot} &= \frac{\nu_{i\cdot}}{n}, \qquad (i = 3, \cdots, r) \\
\hat{p}_{\cdot j} &= \frac{\nu_{\cdot j}}{n}, \qquad (j = 3, \cdots, s).
\end{aligned} \tag{5}$$

Generally if the first r_1 rows ($r_1 < r$) and the first s_1 columns ($s_1 < s$)

are combined with a total of w, we have

$$\hat{p}_{i\cdot} = \frac{\sum_{j=s_1+1}^{s} \nu_{ij} \sum_{i=1}^{r_1} \nu_{i\cdot}}{n \sum_{i=1}^{r_1} \sum_{j=s_1+1}^{s} \nu_{ij}}, \qquad (i = 1, \cdots, r_1)$$

$$\hat{p}_{\cdot j} = \frac{\sum_{i=r_1+1}^{r} \nu_{ij} \sum_{j=1}^{s_1} \nu_{\cdot j}}{n \sum_{i=r_1+1}^{r} \sum_{j=1}^{s_1} \nu_{ij}}, \qquad (j = 1, \cdots, s_1) \tag{6}$$

$$\hat{p}_{i\cdot} = \frac{\nu_{i\cdot}}{n}, \qquad (i = r_1 + 1, \cdots, r)$$

$$\hat{p}_{\cdot j} = \frac{\nu_{\cdot j}}{n}, \qquad (j = s_1 + 1, \cdots, s).$$

The resulting χ^2 has $(r-1)(s-1)-r_1s_1+1$ degrees of freedom.

There is still no difficulty if two pairs of cells, each pair belonging to the same row and both pairs belonging to the same two columns, are combined. Suppose the totals $v=\nu_{11}+\nu_{12}$ and $w=\nu_{21}+\nu_{22}$ are known as well as the separate entries in the remaining cells. We now get

$$\hat{p}_{\cdot 1} = \frac{(\nu_{\cdot 1} + \nu_{\cdot 2})(\nu_{\cdot 1} - \nu_{11} - \nu_{21})}{n(\nu_{\cdot 1} + \nu_{\cdot 2} - v - w)},$$

$$\hat{p}_{\cdot 2} = \frac{(\nu_{\cdot 1} + \nu_{\cdot 2})(\nu_{\cdot 2} - \nu_{12} - \nu_{22})}{n(\nu_{\cdot 1} + \nu_{\cdot 2} - v - w)},$$

$$\hat{p}_{\cdot j} = \frac{\nu_{\cdot j}}{n}, \qquad (j = 3, \cdots, s) \tag{7}$$

$$\hat{p}_{i\cdot} = \frac{\nu_{i\cdot}}{n}, \qquad (i = 1, \cdots, r).$$

The resulting χ^2 has $(r-1)(s-1)-2$ degrees of freedom.

Likewise things work out simply if in each of two different rows a pair of neighboring cells are combined, the two pairs having no column in common. Let $v=\nu_{11}+\nu_{12}$ and $w=\nu_{23}+\nu_{24}$. We find

$$
\begin{aligned}
\hat{p}_{i\cdot} &= \frac{\nu_{i\cdot}}{n}, && (i = 1, \cdots, r)\\
\hat{p}_{\cdot 1} &= \frac{(\nu_{\cdot 1} + \nu_{\cdot 2})(\nu_{\cdot 1} - \nu_{11})}{n(\nu_{\cdot 1} + \nu_{\cdot 2} - v)},\\
\hat{p}_{\cdot 2} &= \frac{(\nu_{\cdot 1} + \nu_{\cdot 2})(\nu_{\cdot 2} - \nu_{12})}{n(\nu_{\cdot 1} + \nu_{\cdot 2} - v)},\\
\hat{p}_{\cdot 3} &= \frac{(\nu_{\cdot 3} + \nu_{\cdot 4})(\nu_{\cdot 3} - \nu_{23})}{n(\nu_{\cdot 3} + \nu_{\cdot 4} - w)},\\
\hat{p}_{\cdot 4} &= \frac{(\nu_{\cdot 3} + \nu_{\cdot 4})(\nu_{\cdot 4} - \nu_{24})}{n(\nu_{\cdot 3} + \nu_{\cdot 4} - w)},\\
\hat{p}_{\cdot j} &= \frac{\nu_{\cdot j}}{n}, && (j = 5, \cdots, s).
\end{aligned} \tag{8}
$$

There is only a slight modification which by now the reader can make for himself in case the two coalesced pairs of adjacent blocks lie in the same row. If $v = \nu_{11} + \nu_{12}$ and $w = \nu_{13} + \nu_{14}$, we have

$$
\begin{aligned}
\hat{p}_{i\cdot} &= \frac{\nu_{i\cdot}}{n}, && (i = 1, \cdots, r)\\
\hat{p}_{\cdot 1} &= \frac{(\nu_{\cdot 1} + \nu_{\cdot 2})(\nu_{\cdot 1} - \nu_{11})}{n(\nu_{\cdot 1} + \nu_{\cdot 2} - v)},\\
\hat{p}_{\cdot 2} &= \frac{(\nu_{\cdot 1} + \nu_{\cdot 2})(\nu_{\cdot 2} - \nu_{12})}{n(\nu_{\cdot 1} + \nu_{\cdot 2} - v)},\\
\hat{p}_{\cdot 3} &= \frac{(\nu_{\cdot 3} + \nu_{\cdot 4})(\nu_{\cdot 3} - \nu_{13})}{n(\nu_{\cdot 3} + \nu_{\cdot 4} - w)},\\
\hat{p}_{\cdot 4} &= \frac{(\nu_{\cdot 3} + \nu_{34})(\nu_{\cdot 4} - \nu_{14})}{n(\nu_{\cdot 3} + \nu_{\cdot 4} - w)},\\
\hat{p}_{\cdot j} &= \frac{\nu_{\cdot j}}{n}, && (j = 5, \cdots, s).
\end{aligned} \tag{9}
$$

The degrees of freedom for χ^2 in both the last two cases is $(r-1)(s-1)-2$.

There is no difficulty so long as the pairs or rectangular blocks of neighboring cells have all of their rows or columns in common or else have no rows or columns in common. As soon as there is partial over-

lapping of rows or columns, the likelihood equations are not so simply solved. Consider the simplest case of this kind in which $v = \nu_{11} + \nu_{12}$ and $w = \nu_{22} + \nu_{23}$. The likelihood (or modified χ^2-minimum) equations for rows are as simple as ever, giving

$$\hat{p}_{i\cdot} = \frac{\nu_{i\cdot}}{n}, \qquad (i = 1, 2, \cdots, r).$$

But the equations for columns are

$$\begin{aligned}
\frac{v}{p_{\cdot 1} + p_{\cdot 2}} + \frac{\nu_{\cdot 1} - \nu_{11}}{p_{\cdot 1}} &= \frac{\nu_{\cdot s}}{p_{\cdot s}}, \\
\frac{v}{p_{\cdot 1} + p_{\cdot 2}} + \frac{w}{p_{\cdot 2} + p_{\cdot 3}} + \frac{\nu_{\cdot 2} - \nu_{12} - \nu_{22}}{p_{\cdot 2}} &= \frac{\nu_{\cdot s}}{p_{\cdot s}}, \\
\frac{w}{p_{\cdot 2} + p_{\cdot 3}} + \frac{\nu_{\cdot 3} - \nu_{23}}{p_{\cdot 3}} &= \frac{\nu_{\cdot s}}{p_{\cdot s}}, \\
\frac{\nu_{\cdot 4}}{p_{\cdot 4}} &= \frac{\nu_{\cdot s}}{p_{\cdot s}}, \\
&\vdots \\
\frac{\nu_{\cdot (s-1)}}{p_{\cdot (s-1)}} &= \frac{\nu_{\cdot s}}{p_{\cdot s}}.
\end{aligned} \tag{10}$$

On multiplying by corresponding $p_{\cdot j}$'s and adding, we still get

$$\hat{p}_{\cdot j} = \frac{\nu_{\cdot j}}{n}, \qquad (j = 4, \cdots, s).$$

To solve the remaining three equations set

$$\frac{p_{\cdot 1}}{p_{\cdot 1} + p_{\cdot 2} + p_{\cdot 3}} = a, \qquad \frac{p_{\cdot 2}}{p_{\cdot 1} + p_{\cdot 2} + p_{\cdot 3}} = b, \qquad c = 1 - a - b;$$

$\nu_{\cdot 1} + \nu_{\cdot 2} + \nu_{\cdot 3} = n_1$; and

$$\nu_{\cdot 1} - \nu_{11} = A_1, \qquad \nu_{\cdot 2} - \nu_{12} - \nu_{22} = A_2, \qquad \nu_{\cdot 3} - \nu_{23} = A_3.$$

Then on eliminating a from the first and third equations by means of

$$a = \frac{A_1(1 - c)}{n_1(1 - c) - v}, \tag{11}$$

we have the quadratic for c:

$$n_1(n_1 - w - A_1)c^2 - [(n_1 - v)(n_1 - w) - n_1(A_1 - A_3) - A_1A_3]c + (n_1 - v - A_1)A_3 = 0. \tag{12}$$

It can be shown by some fairly messy but elementary algebra that the two roots of this equation are both real and are positive proper fractions. Further the larger root cannot be used since then, using (11), we get $a+c>1$. Thus, taking the smaller root of (12), we get a unique set of values for a, b, and c which uniquely determine $p_{\cdot 1}$, $p_{\cdot 2}$, and $p_{\cdot 3}$. The degrees of freedom for χ^2 are still $(r-1)(s-1)-2$.

REFERENCES

[1] Freeman, G. H., and Halton, J. H., "Note on an exact treatment of contingency, goodness of fit and other problems of significance," *Biometrika*, 38 (1951), 141–9.

[2] Adler, Franz, "Yates' correction and the statisticians," *Journal of the American Statistical Association*, 46 (1951), 490–501.

[3] Cochran, W. G., "The χ^2 test of goodness of fit," *Annals of Mathematical Statistics*, 23 (1952), 315–45.

[4] Cramér, Harald, *Mathematical Methods of Statistics*. Princeton: Princeton University Press, 1946. Sections 30.3 and 30.5, pp. 424–34 and 441–45.

[5] Hoel, P. G., *Introduction to Mathematical Statistics*. New York: John Wiley & Sons, 1947, p. 192.

APPROXIMATING THE MODE FROM WEIGHTED SAMPLE VALUES

HOWARD L. JONES
Illinois Bell Telephone Company

The weighted mean of ordered sample observations can be used to approximate the mode under favorable conditions, where the weights are determined from the first two terms in a Taylor expansion of the maximum likelihood estimate. Such weights are shown in Table 1 for the case where the sample is selected from a *t*-distribution with known kurtosis.

INTRODUCTION

ONE of the oldest and most difficult statistical problems is that of deciding how much weight to assign to extreme sample observations in trying to estimate the typical or average measurement for some population. It may be regarded as a part of the more general problem of how to make estimates on the basis of observed sample values after such values have been arranged in order of size.

Sample values thus arranged are now commonly called *order statistics;* and estimates computed therefrom, as well as other functions of these statistics, have been termed *systematic statistics* [4]. Weighted sums of sample values arranged in order of size may therefore be called *linear systematic statistics* [3]. Examples are the sample median, where a weight of zero is assigned to each order statistic except for one or two central values (depending on whether the sample size is odd or even), and the midrange where the first and last order statistic is each weighted one-half and every other order statistic is weighted zero. The sample mean may also be regarded as a systematic statistic where each order statistic is assigned a weight equal to the reciprocal of the sample size; although in this case, the equality of the weights makes it immaterial whether the sample values are ordered or not.

The advantages of linear systematic statistics have been discussed by various writers. It is known [1, p. 483] that the sample mean is the best unbiased regular estimate of the mean of a normal distribution. It also tends to have desirable asymptotic properties in large samples from most types of distributions that statisticians have studied. But even where other types of functions are more efficient in the sense that for a given sample size they lead to unbiased estimates with a smaller variance, a linear systematic statistic may have important advantages in practical problems. Chief among these are the simplicity of the computations, and the economy of time and expense in situations where

samples of any desired size can be selected and measured with relative ease [2]. These advantages explain why the range (a systematic statistic in which the extreme order statistics are weighted -1 and $+1$), instead of the root-mean-square, is generally used to estimate the standard deviation of a distribution in quality control applications. Some work has been done on problems of measuring and maximizing the efficiency of various linear statistics, particularly for samples from a normal distribution [5].

A deterrent to the use of these statistics is the difficulty of finding satisfactory weights. The general problem was discussed several years ago in an unpublished paper, "The Weighted Mean of Random Observations Arranged in Order of Size," which was written jointly with John H. Smith and recently reproduced in Ditto form as National Bureau of Standards Working Paper SEL-51-1, February 1951. The weights which lead to the best unbiased linear estimates of the mean, mode, median, or any other location parameter can usually be determined when the first and second moments and the product moments of all the order statistics are known. Since these moments are difficult to evaluate for many typical distributions, a general method was given for finding weights that would lead to a first approximation of the mode under certain favorable conditions. The present article is the result of Churchill Eisenhart's desire to have this method published more widely, since it has been found to be useful in broader applications.

In order to illustrate an application of the method, the weights shown in Table 1 were computed for various sized samples selected from a t-distribution with varying degrees of kurtosis. Hence, if one knows that the parent distribution is approximately of the same shape as some form of the t-distribution, and if the approximate kurtosis of this distribution is also known, the mode can be estimated from a sample of size 10 or less by arranging the sample values in order of size and then applying the weights shown in the table. If the kurtosis is 3, indicating a normal distribution, the weights are all equal, as they should be. If the kurtosis is large, the median is weighted more heavily than the extreme observations.

While the mode of a distribution is generally of less interest than the mean or even the median, there are many distributions like the t-distributions where these location parameters coincide. The estimated location of the mode may also be useful in other situations where its relationship to other parameters like the mean or median is known. The mode is emphasized here merely because it is important in the proposed method for developing the weights.

TABLE 1

WEIGHTS FOR APPROXIMATING THE MODE OF A *t*-DISTRIBUTION

Size of Sample	Rank in Sample	Kurtosis:* $\alpha_4 = \mu_4/\mu_2^2$							
		3	3.5	4	4.5	5	6	9	∞
3	2	.34	.36	.36	.38	.38	.40	.40	.42
	1, 3	.33	.32	.32	.31	.31	.30	.30	.29
4	2, 3	.25	.26	.28	.28	.29	.29	.30	.31
	1, 4	.25	.24	.22	.22	.21	.21	.20	.19
5	3	.20	.22	.22	.24	.24	.24	.26	.28
	2, 4	.20	.21	.22	.22	.22	.23	.23	.24
	1, 5	.20	.18	.17	.16	.16	.15	.14	.12
	3, 4	.17	.18	.19	.20	.20	.21	.22	.23
6	2, 5	.17	.17	.18	.18	.18	.18	.18	.18
	1, 6	.16	.15	.13	.12	.12	.11	.10	.09
	4	.14	.16	.16	.18	.18	.18	.20	.20
	3, 5	.15	.16	.16	.17	.17	.18	.18	.19
7	2, 6	.14	.14	.15	.14	.15	.14	.14	.14
	1, 7	.14	.12	.11	.10	.09	.09	.08	.07
	4, 5	.13	.14	.15	.15	.16	.16	.17	.18
8	3, 6	.13	.14	.14	.15	.15	.15	.15	.16
	2, 7	.12	.12	.12	.12	.12	.12	.12	.11
	1, 8	.12	.10	.09	.08	.07	.07	.06	.05
	5	.12	.12	.14	.14	.14	.14	.16	.16
9	4, 6	.11	.12	.13	.14	.14	.14	.15	.16
	3, 7	.11	.12	.12	.12	.13	.13	.13	.13
	2, 8	.11	.11	.11	.10	.10	.10	.10	.09
	1, 9	.11	.09	.07	.07	.06	.06	.04	.04
	5, 6	.10	.11	.12	.12	.13	.13	.14	.15
	4, 7	.10	.11	.12	.12	.12	.13	.13	.14
10	3, 8	.10	.10	.11	.11	.11	.11	.11	.11
	2, 9	.10	.10	.09	.09	.09	.08	.08	.07
	1, 10	.10	.08	.06	.06	.05	.05	.04	.03

* The column headings are the values of α_4 for the parent distribution, not for the sample.

DERIVATION OF THE WEIGHTS

For some infinite population, assume that the distribution of a measurement x is described by the probability density function

$$y = f(x - \theta_1, \theta_2, \theta_3, \cdots, \theta_k), \tag{1}$$

where $\theta_1, \theta_2, \cdots, \theta_k$ are parameters of the distribution, the form of the function being such that θ_1 is a location parameter. Assume, for every admissible value of x where $y>0$, that the derivative $y'=dy/dx$ exists, that y has a unique mode at $x=\theta_1$, and that y'/y is analytic. Let y_i and y_i' respectively, denote the value of y and its first derivative at the point $x=x_i$; and let

$$\phi(x_i) = -y_i'/y_i, \tag{2}$$

$$\phi'(x_i) = d\phi(x_i)/dx_i. \tag{3}$$

Assume that for every pair x_i and x_j, the ratio

$$R_{ij} = \phi'(x_i)/\phi'(x_j) \tag{4}$$

for $\theta_1=0$ is known; that is, the value of R_{ij} does not involve an unknown parameter. Then weights for approximating the mode can be developed as follows.

Let

$$y_r = f(x_r - \theta_1, \theta_2, \cdots, \theta_k) \tag{5}$$

where x_r is the rth order statistic in a random sample of size n from distribution (1). The maximum likelihood estimate of θ_1 when $\theta_2, \theta_3, \cdots, \theta_k$ are known is, in general, a solution of the equation

$$\frac{\partial}{\partial\theta_1} \sum \log f(x_r - \theta_1, \theta_2, \cdots, \theta_k) = 0, \qquad r = 1, 2, \cdots, n, \tag{6}$$

where $\sum$ denotes summation over all n values of r. From the juxtapositional relationship between x and θ_1 in (5),

$$\frac{\partial}{\partial\theta_1} \log f(x_r - \theta_1, \theta_2, \cdots, \theta_k) = -\frac{d}{dx_r} \log f(x_r - \theta_1, \theta_2, \cdots, \theta_k) = \phi(x_r), \tag{7}$$

where $\phi(x_r)$ is defined by (2). Combining (6) and (7), we obtain

$$\sum \phi(x_r) = 0. \tag{8}$$

Let $\hat{\theta}_1$ be a value of θ_1 that satisfies (8). Expanding $\phi(\hat{\theta}_1)$ in the neighborhood of x_r yields

$$\phi(\hat{\theta}_1) = \phi(x_r) + (\hat{\theta}_1 - x_r)\phi'(x_r) + \tfrac{1}{2}(\hat{\theta}_1 - x_r)^2\phi'(x_r) + \cdots, \tag{9}$$

where $\phi'(x_r)$, $\phi''(x_r)$, . . . denote successive derivatives with respect to x_r. Neglecting all but the first two terms of the expansion (which procedure may introduce considerable error when $\hat{\theta}_1 - x_r$ is large), and summing over the n observations, we obtain the approximate relationship,

$$n\phi(\hat{\theta}_1) \sim \sum \phi(x_r) + \sum (\hat{\theta}_1 - x_r)\phi''(x_r), \tag{10}$$

or after combining with (8),

$$n\phi(\hat{\theta}_1) \sim \sum (\hat{\theta}_1 - x_r)\phi'(x_r). \tag{11}$$

Now suppose that $\hat{\theta}_1 = \theta_1$, the mode of distribution (1). (In general, this supposition will be true only to a first approximation.) Then from (2),

$$\phi(\hat{\theta}_1) = 0. \tag{12}$$

Combining (11) and (12) leads to

$$\hat{\theta}_1 = \sum w_r x_r \tag{13}$$

where

$$w_r = \phi'(x_r) / \sum \phi'(x_r). \tag{14}$$

For a normal distribution, $\phi'(x_r) = 1/\sigma^2$ for every x_r; hence, $w_r = 1/n$ for this case, which is as it should be. In general, however, $\phi'(x)$ varies with x; and in order to make use of (14) to compute a linear estimate of the mode, we must choose some value of $\phi'(x_r)$—say $\phi'(\ddot{x}_r)$—that is typical for the order statistic x_r. In most applications, perhaps the simplest choice is the one where $\ddot{x}_r$ is the mode of the distribution of x_r in a sample of size n from a parent distribution of the specified form with the mode chosen as the origin so that θ_1 does not appear in the general expression for w_r. This value of x_r can usually be determined or approximated when the form of the parent distribution is known.

Let $F(x)$ denote the cumulative probability distribution of x (that is, the probability that a random selection from distribution (1) will be equal to or less than a particular value of x). Then the distribution of the order statistic x_r (where $x_r \leqq x_{r+1}$) may be written as

$$\frac{n!}{(r-1)!(n-r)!} [F(x_r)]^{r-1}[1 - F(x_r)]^{n-r} y_r. \tag{15}$$

To find $\ddot{x}_r$, the mode of this distribution, we set the derivative equal to zero. After dividing by

$$\frac{n!}{(r-1)!(n-r)!}[F(x_r)]^{r-1}[1-F(x_r)]^{n-r}y_r^2,$$

we obtain the equation

$$\frac{r-1}{F(x_r)} - \frac{n-r}{1-F(x_r)} + \frac{y_r'}{y_r^2} = 0. \tag{16}$$

This equation can generally be solved for x_r, for $r=1, 2, \cdots, n$, by an iterative procedure.

COMPUTING THE WEIGHTS FOR THE t-DISTRIBUTION

The probability density function for the t-distribution may be written in the form

$$y = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}}\left[1 + \frac{(t-\theta)^2}{k}\right]^{-(k+1)/2} \tag{17}$$

where k is the number of degrees of freedom. Since the mode is at $t=\theta$, we set $\theta=0$ and obtain the functions

$$y_r' = -\frac{(k+1)t_r}{k+t_r^2}y_r, \tag{18}$$

$$\phi(t_r) = \frac{(k+1)t_r}{k+t_r^2}, \tag{19}$$

$$\phi'(t_r) = \frac{(k+1)(k-t_r^2)}{(k+t_r^2)^2}. \tag{20}$$

From (14) and (20), the desired weights are seen to be proportional to $(k-\check{t}_r^2)/(k+\check{t}_r^2)^2$, where $\check{t}_r$ is a typical value of t_r.

Let us first consider the case where $k=1$. After dividing by π, equation (16) for this case becomes equal to

$$\frac{r-1}{\frac{1}{2}\pi + \arctan t_r} - \frac{n-r}{\frac{1}{2}\pi - \arctan t_r} - 2t_r = 0, \tag{21}$$

where the angles are measured in radians. For $n=5$, solving (21) for $1, 2, \cdots, 5$ successively by an iterative procedure and employing

(20) and (14) to find the corresponding weights, we obtain the following approximations:

r	$\check{t}_r$	$\phi'(\check{t}_r)$	w_r
1	−.8738	.1521	.03
2	−.3689	1.3387	.27
3	.0000	2.0000	.40
4	.3689	1.3387	.27
5	.8738	.1521	.03

The weights shown in the last column indicate that in a sample consisting of five observations from this distribution, the sample median should receive the greatest weight (.40) and the two extreme terms very little weight (.03) in estimating the mode.

Suppose we make the transformation

$$k = 2p, \qquad t = \sqrt{2p(1-x)/x}, \tag{22}$$

for $t>0$. Then for $x=x_r$, (20) reduces to

$$\phi'(x_r) = \frac{2p+1}{2p}\, x_r(2x_r - 1). \tag{23}$$

The weights should therefore be proportional to $\check{x}_r(2\check{x}_r-1)$, where $\check{x}_r$ is the value of x corresponding to the mode of the distribution of t_r. Since p and x have the same meaning here as in Pearson's Tables of the Incomplete Beta Function, we write

$$F(t_r) = 1 - \tfrac{1}{2}I(p, .5), \tag{24}$$

$$\frac{1}{y_r^2}\,\frac{dy_r}{dt_r} = -(2p+1)\,\frac{\sqrt{1-x_r}}{x_r^p}\,B(p, .5), \tag{25}$$

where $I(p, .5)$ and $B(p, .5)$ also have the same meaning as in Pearson's tables. Equation (16) can therefore be written as

$$\frac{r-1}{1-\tfrac{1}{2}I(p, .5)} - \frac{n-r}{\tfrac{1}{2}I(p, .5)} - (2p+1)\,\frac{\sqrt{1-x_r}}{x_r^p}\,B(p, .5) = 0, \tag{26}$$

and approximate solutions for $r=1, 2, \cdots > \frac{1}{2}n$ can be found by interpolating in those tables. These solutions can be taken as the values of x_r to be substituted in equation (23) in computing $\phi'(x_r)$ and w_r.

The procedure just described was used in arriving at the weights shown in Table 1. The kurtosis shown in the column headings is related to the number of degrees of freedom by the equation,

$$\alpha_4 = 3(k-2)/(k-4), \qquad \text{for } k \geqq 4. \tag{27}$$

COMPUTING THE WEIGHTS FOR A POLYNOMIAL DISTRIBUTION

The properties of the estimates discussed in this article have not been investigated for the general case. Investigation for cases where the distribution is a polynomial in x is relatively easy, however, provided the sample size and the degree of the polynomial are small. The following example shows what happens in a particular situation of this kind where the distribution is skewed.

Consider the distribution of the simple form

$$z = 12u^2(1-u), \qquad 0 \leqq u \leqq 1, \tag{28}$$

which is a special case for the derivative of the Incomplete Beta Function. We first note that the mode is at $u = \frac{2}{3}$. Distributions of this same general form may therefore be obtained by making the transformation

$$u = \frac{x-\theta_1}{\theta_2} + \frac{2}{3}. \tag{29}$$

Since $du/dx = 1/\theta_2$, equation (28) becomes

$$y = \frac{12}{\theta_2}\left(\frac{x-\theta_1}{\theta_2} + \frac{2}{3}\right)^2\left(\frac{1}{3} - \frac{x-\theta_1}{\theta_2}\right), \tag{30}$$

with x satisfying the inequality

$$\theta_1 - \frac{2\theta_2}{3} \leqq x \leqq \theta_1 + \frac{\theta_2}{3}. \tag{31}$$

This distribution has its mode at $x = \theta_1$, and is of the form prescribed in discussing distribution (1).

After setting $\theta_1 = 0$ and $\theta_2 = 1$, we may proceed as previously suggested and write down the relationships[1]

$$y_r = \frac{4}{9}(1-3x_r)(2+3x_r)^2, \tag{32}$$

[1] The computations can be carried out without assigning the arbitrary value of 1 to the scale parameter, θ_1. The effect will be to multiply the resulting values of $\phi'(\tilde{x}_r)$ by the reciprocal of θ_2^3.

$$y_r' = -12x_r(2 + 3x_r), \tag{33}$$

$$\phi(x_r) = -y_r'/y_r = \frac{27x_r}{(1 - 3x_r)(2 + 3x_r)}, \tag{34}$$

$$\phi'(x_r) = \frac{27(2 + 9x_r^2)}{(1 - 3x_r)^2(2 + 3x_r)^2}. \tag{35}$$

From (14), the desired weights are to be proportional to (35) for $r = 1, 2, \cdots, n$, where x_r is a typical value of the rth order statistic.

Suppose the sample is of size 3. Employing (16), we may find the modal values of x_1, x_2, and x_3 from the relationship

$$\frac{r - 1}{F(\breve{x}_r)} - \frac{3 - r}{1 - F(\breve{x}_r)} - \frac{243\breve{x}_r}{4(1 - 3\breve{x}_r)^2(2 + 3\breve{x}_r)^3} = 0, \tag{36}$$

where

$$-\frac{2}{3} \leqq \breve{x}_r \leqq \frac{1}{3}, \tag{37}$$

and

$$F(\breve{x}_r) = \int_{-2/3}^{\breve{x}_r} y_r dx_r = \frac{1}{27}(2 - 3\breve{x}_r)(2 + 3\breve{x}_r)^3. \tag{38}$$

We may also write

$$1 - F(\breve{x}_r) = \frac{1}{27}(1 - 3\breve{x}_r)^2(11 + 18\breve{x}_r + 9\breve{x}_r^2). \tag{39}$$

Putting $r = 1$, 2, and 3 successively in equation (36) and clearing of fractions, we obtain the equations

$$297\breve{x}_1^3 + 594\breve{x}_1^2 + 387\breve{x}_1 + 64 = 0, \tag{40}$$

$$891\breve{x}_2^4 + 1188\breve{x}_2^3 - 27\breve{x}_2^2 - 582\breve{x}_2 - 20 = 0, \tag{41}$$

$$99\breve{x}_3^2 - 66\breve{x}_3 + 8 = 0. \tag{42}$$

The roots of these equations satisfying (37) have the approximate values

$$\breve{x}_1 = -.2481851535, \tag{43}$$

$$\breve{x}_2 = -.0345011424, \tag{44}$$

$$\breve{x}_3 = .1592556774. \tag{45}$$

Substituting these values for x_r in (35), we obtain the approximations

$$\phi'(\dot{x}_1) = 14.37744535, \tag{46}$$

$$\phi'(\dot{x}_2) = 12.39545132, \tag{47}$$

$$\phi'(\dot{x}_3) = 35.9319163. \tag{48}$$

Using these values in (14), we find the approximate values of the weights to be as follows:

$$w_1 = .229287748, \tag{49}$$

$$w_2 = .197679424, \tag{50}$$

$$w_3 = .573032828. \tag{51}$$

In using these weights, we are likely to be especially interested in three properties of the resulting estimates of the mode, these properties being defined by the following relationships:

$$\text{Bias} = E[\hat{\theta}_1] - \theta_1, \tag{52}$$

$$\text{Variance} = E[\hat{\theta}_1^2] - (E[\hat{\theta}_1])^2, \tag{53}$$

$$\begin{aligned} \text{Mean Square Error} &= E[(\hat{\theta}_1 - \theta_1)^2] \\ &= (\text{Bias})^2 + \text{Variance}, \end{aligned} \tag{54}$$

where the symbol E denotes average or expected value. For this particular case where the distribution is described by equation (30), these properties when the weights in (49), (50), and (51) are used to estimate the mode are indicated in Table 2 in the column headed "'Maximum likelihood' estimates." For purposes of comparison, the same properties are shown for two other weight systems, the second system being that which yields the "best unbiased estimates"—that is, the unbiased linear estimates with minimum variance—and the third being the system which yields the linear estimates with minimum mean square error.[2] The methods of deriving these weight systems and their properties are outlined in the appendix that follows.

It should be kept in mind that the comparison shown in Table 2 applies only to distributions of the form described by equation (30), and

[2] It is assumed here that attention is restricted to linear estimates for which the sum of the weights is equal to one, and that there is no *a priori* information about the size of the parameters (aside from the implied specification that θ_1 is positive). Under other conditions, the weights shown in the last two columns of Table 2 may not be optimum with respect to the properties indicated in the column headings. For example, if θ_1 is known to satisfy the inequality

$$-.000000001 < \theta_1 < .000000001$$

and θ_2 is known to be large, say

$$\theta_2 > 100{,}000,$$

then the estimates obtained by employing the weights shown in the last column of the table would not be as good as estimates obtained by setting all the weights equal to zero.

may not be indicative of the relative merits of different methods of estimating the mode for other types of distribution.

TABLE 2

COMPARISON OF THREE LINEAR SYSTEMATIC ESTIMATES OF THE MODE OF A POLYNOMIAL DISTRIBUTION*

	"Maximum likelihood" estimates	"Best" unbiased estimate	Minimum mean square error estimate
Weights			
w_1	.229287748	.232616040	.241695273
w_2	.197679424	.133288817	.134304834
w_3	.573032828	.634095143	.623999893
Bias	$-.0092025370_2$	0	$-.0032760920_2$
Variance	$.0124648780_2^2$	$.0124789160_2^2$	$.0124549140_2^2$
Mean square error	$.0125495640_2^2$	$.0124789160_2^2$	$.0124656470_2^2$

* The values shown are for the particular polynomial distribution described by equation (30). The variance of this distribution is $.040_1^3$.

APPENDIX

Lower moments of the order statistics

To investigate the properties of various linear systematic estimates of the mode of the distribution described by equation (30), we first derive the lower moments of the order statistics. Consider the distribution described by equation (28). For a sample of size three from this distribution, the moments and product moments of the order statistics may be obtained from the general formula

$$(55)\quad E[u_1^h u_2^i u_3^k] = 3! \int_{u_3=0}^{1} \int_{u_2=0}^{u_3} \int_{u_1=0}^{u_2} u_1^h u_2^i u_3^k y_1 y_2 y_3 du_1 du_2 du_3$$

where

$$(56)\qquad y_i = 12u_i^2(1 - u_i), \qquad i = 1, 2, 3.$$

It turns out that

$$(57)\qquad E[u_1] = \frac{2127}{5005}, \qquad E[u_2] = \frac{3039}{5005}, \qquad E[u_3] = \frac{3843}{5005},$$

$$(58)\qquad E[u_1^2] = \frac{1028}{5005}, \qquad E[u_2^2] = \frac{1948}{5005}, \qquad E[u_3^2] = \frac{3030}{5005}.$$

$$(59)\qquad E[u_1u_2] = \frac{6777}{25025}, \qquad E[u_1u_3] = \frac{333}{1001}, \qquad E[u_2u_3] = \frac{477}{1001}.$$

The lower moments of order statistics for samples of size three from the distribution described by equation (30) can now be obtained by direct substitution of the right side of (29) for u in (57), (58), and (59), since there is a linear relationship between u and x. Thus, we obtain

$$E[x_1] = \theta_1 - \frac{3629}{15015}\theta_2, \tag{60}$$

$$E[x_2] = \theta_1 - \frac{893}{15015}\theta_2,$$

$$E[x_3] = \theta_1 + \frac{1519}{15015}\theta_2,$$

$$E[x_1^2] = \theta_1^2 - \frac{7258}{15015}\theta_1\theta_2 + \frac{3748}{45045}\theta_2^2, \tag{61}$$

$$E[x_2^2] = \theta_1^2 - \frac{1786}{15015}\theta_1\theta_2 + \frac{1084}{45045}\theta_2^2,$$

$$E[x_3^2] = \theta_1^2 + \frac{3038}{15015}\theta_1\theta_2 + \frac{1174}{45045}\theta_2^2,$$

$$E[x_1x_2] = \theta_1^2 - \frac{4522}{15015}\theta_1\theta_2 + \frac{6113}{225225}\theta_2^2, \tag{62}$$

$$E[x_1x_3] = \theta_1^2 - \frac{422}{3003}\theta_1\theta_2 - \frac{163}{9009}\theta_2^2,$$

$$E[x_2x_3] = \theta_1^2 + \frac{626}{15015}\theta_1\theta_2 + \frac{193}{45045}\theta_2^2.$$

From these values, the variances and covariances of the order statistics are found to be

$$\sigma_1^2 = E[x_1^2] - (E[x_1])^2 = \frac{621,011}{25,050,025}\theta_2^2, \tag{63}$$

$$\sigma_2^2 = E[x_2^2] - (E[x_2])^2 = \frac{514,219}{25,050,025}\theta_2^2,$$

$$\sigma_3^2 = E[x_3^2] - (E[x_3])^2 = \frac{396,501}{25,050,025}\theta_2^2,$$

$$\sigma_{12} = E[x_1x_2] - E[x_1]E[x_2] = \frac{319,824}{25,050,025}\,\theta_2^2, \tag{64}$$

$$\sigma_{13} = E[x_1x_3] - E[x_1]E[x_3] = \frac{159,264}{25,050,025}\,\theta_2^2,$$

$$\sigma_{23} = E[x_2x_3] - E[x_2]E[x_3] = \frac{258,048}{25,050,025}\,\theta_2^2.$$

Lower moments of linear estimates of the mode

A linear function

$$\hat{\theta}_1 = w_1x_1 + w_2x_2 + w_3x_3 \tag{65}$$

of the order statistics x_1, x_2, x_3 in a sample of size three from the distribution described by equation (30) has the mean value

$$E[\hat{\theta}_1] = w_1E[x_1] + w_2E[x_2] + w_3E[x_3]. \tag{66}$$

Its variance is

$$E[\hat{\theta}_1^2] - (E[\hat{\theta}_1])^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + w_3^2\sigma_3^2 + 2(w_1w_2\sigma_{12} + w_1w_3\sigma_{13} + w_2w_3\sigma_{23}). \tag{67}$$

Combining (66) with (60), and (67) with (63) and (64), and employing the weights in (49), (50), and (51), we obtain

$$E[\hat{\theta}_1] = \theta_1 - .009202537\theta_2, \tag{68}$$

$$E[\hat{\theta}_1^2] - (E[\hat{\theta}_1])^2 = .012464878\theta_2^2. \tag{69}$$

The bias, obtained by subtracting θ_1 from (68), is therefore $-.009202537\theta_2$, while the variance is given by (69). Adding the square of the bias to the variance, we obtain

$$E[(\hat{\theta}_1 - \theta_1)^2] = .012549564\theta_2^2, \tag{70}$$

which is the mean square error of the estimates. These results are shown in Table 2 in the column headed " 'Maximum likelihood' estimates."

To yield unbiased estimates of the mode of distribution (30) for all values of θ_1 and θ_2, the weights w_1, w_2, w_3 must satisfy the condition

$$w_1E[x_1] + w_2E[x_2] + w_3E[x_3] = \theta_1 \tag{71}$$

for every θ_1 and θ_2. Combining (71) with (60) leads to

$$w_1\left(\theta_1 - \frac{3629}{15015}\theta_2\right) + w_2\left(\theta_1 - \frac{893}{15015}\theta_2\right) + w_3\left(\theta_1 + \frac{1519}{16015}\theta_2\right) = \theta_1, \tag{72}$$

which must be satisfied identically. Hence, the weights must satisfy the two conditions

$$w_1 + w_2 + w_3 = 1, \tag{73}$$

$$3629w_1 + 893w_2 - 1519w_3 = 0. \tag{74}$$

Combining (67) with (63) and (64), the variance of any linear function of the order statistics x_1, x_2, and x_3 is found to be

$$\frac{\theta_2^2}{25{,}050{,}025}\{621{,}011w_1^2 + 514{,}219w_2^2 + 396{,}501w_3^2 + 2(319{,}824w_1w_2 + 159{,}264w_1w_3) + 258{,}048w_2w_3)\}. \tag{75}$$

To find the regular unbiased linear estimate with minimum variance, we write down the expression

$$621{,}011w_1^2 + 514{,}219w_2^2 + 396{,}501w_3^2 + 2(319{,}824w_1w_2 + 159{,}264w_1w_3 + 258{,}048w_2w_3) + \lambda_1(w_1 + w_2 + w_3 - 1) + \lambda_2(3629w_1 + 893w_2 - 1519w_3), \tag{76}$$

where λ_1 and λ_2 are Lagrange multipliers, and set its partial derivatives with respect to w_1, w_2, w_3 each equal to zero. After eliminating λ_1 and λ_2, we obtain the equation

$$613w_1 - 2497w_2 + 300w_3 = 0. \tag{77}$$

Solving (73), (74), and (77) simultaneously for w_1, w_2, and w_3 yields

$$w_1 = 3{,}525{,}043/15{,}153{,}912, \tag{78}$$

$$w_2 = 2{,}019{,}847/15{,}153{,}912, \tag{79}$$

$$w_3 = 9{,}609{,}022/15{,}153{,}912. \tag{80}$$

The variance and mean square error of the estimates computed with these weights is shown in Table 2 in the column headed "'Best' unbiased estimates."

The mean square error of the estimates considered here may be found directly by setting $\theta_1 = 0$ in equations (61) and (62) and then computing the value of the expression

$$w_1^2E[x_1^2] + w_2^2E[x_2^2] + w_3^2E[x_3^2] + 2(w_1w_2E[x_1x_2] + w_1w_3E[x_1x_3] + w_2w_3E[x_2x_3]), \tag{81}$$

which is now equivalent to

$$\frac{\theta_2^2}{225,225}(18,740w_1^2 + 5,420w_2^2 + 5,870w_3^2 + 12,226w_1w_2 - 8,150w_1w_3 + 1,930w_2w_3). \tag{82}$$

The regular linear estimate with minimum mean square error may be found by writing down the expression

$$18,740w_1^2 + 5,420w_2^2 + 5,870w_3^2 + 12,226w_1w_2 - 8,150w_1w_3 + 1,930w_2w_3 + \lambda(w_1 + w_2 + w_3 - 1), \tag{83}$$

and then setting the partial derivatives with respect to w_1, w_2, and w_3 each equal to zero. After eliminating λ, we obtain the equations

$$12,627w_1 + 693w_2 - 5,040w_3 = 0, \tag{84}$$

$$10,188w_1 + 4,455w_2 - 4,905w_3 = 0. \tag{85}$$

Solving these two equations and (73) simultaneously leads to

$$w_1 = 18,095/74,867, \tag{86}$$

$$w_2 = 10,055/74,867, \tag{87}$$

$$w_3 = 46,717/74,867. \tag{88}$$

The bias, variance, and mean square error of the estimates obtained from these weights are shown in the last column of Table 2. The weights for all three types of estimates covered by the table are shown there in decimal form to permit ready comparison.

REFERENCES

[1] Cramér, H., *Mathematical Methods of Statistics*. Princeton: Princeton University Press, 1946.

[2] Dixon, W. J., and Massey, F. J., Jr., *Introduction to Statistical Analysis*. New York: McGraw-Hill Book Co., 1951.

[3] Godwin, H. J., "On the estimation of dispersion by linear systematic statistics," *Biometrika*, 36 (1949), 92–100.

[4] Mosteller, F., "On some useful 'inefficient' statistics," *Annals of Mathematical Statistics*, 17 (1946), 377–408.

[5] Nair, K. R., "Efficiencies of certain linear systematic statistics for estimating dispersion from normal samples," *Biometrika*, 37 (1950), 182–83.

ESTIMATING THE RATIO BETWEEN THE PROPORTIONS OF TWO CLASSES WHEN ONE IS A SUB-CLASS OF THE OTHER

JACK M. ELKIN
U. S. Railroad Retirement Board

1. Consider a population of size N, of which p_1N items have the characteristic A, of which, in turn, p_2N also have the characteristic B; i.e., the proportion of A-items which are also B-items is p_2/p_1. For example, p_1 may be the proportion of individuals eligible to retire among a group of individuals covered by a retirement program at the beginning of the year and p_2/p_1 the proportion of eligibles who retire during the year, or p_1 may be the proportion of families in a given income class among a group of families in all income classes and p_2/p_1 the proportion of families in the specified income class who own their own homes. A straightforward problem is one of estimating p_2/p_1 from a simple random sample of size n drawn from N. The problem as stated in the title is, of course, equivalent to one of estimating the ratio of the overlap between two overlapping classes to one of the classes.

Sometimes the actual value of p_1 or p_2 is known beforehand. On the "theory" that it is always better to use a known than an estimated datum, one might be tempted to estimate p_2/p_1 by estimating the unknown proportion from the sample and combining that estimate with the known proportion. This, however, is not necessarily the best course to follow. If we consider the "better" of two estimates to be the one with the smaller coefficient of variation, it can be shown, for a sufficiently large sample, that

(a) even if p_2 is known, a better estimate of p_2/p_1 (or of p_1/p_2) can be obtained by estimating both proportions from the sample if $p_2 > p_1/(2-p_1)$, or, alternatively, if $p_1 < 2p_2/(1+p_2)$, and

(b) even if p_1 is known, a better estimate of p_2/p_1 (or of p_1/p_2) can be obtained by estimating both proportions from the sample, regardless of the values of p_1 and p_2.

The criterion in (a) cannot be applied exactly in practice, since p_1 is not known, but its sample value can almost always be satisfactorily substituted for the universe value in the inequalities.

By way of illustration, suppose that, on the basis of a simple random sample taken of 100,000 employees covered by a pension program at the beginning of a year, it is estimated that 5,000 were eligible to retire during the year and that 3,000 of these did retire. Suppose, further, that the exact number of those who retired is known, from another

source, to be 3,200. Since $.03 > .05/(2-.05)$, the number of retirements as a proportion of the number eligible for retirement should be quoted as $\frac{3}{5}$ or 60 per cent, rather than 3.2/5 or 64 per cent. On the other hand, suppose that the sample estimate of the number who retired is 2,000, but that the exact number is known to be 2,100. Since $.02 < .05/(2-.05)$, the proportion should be quoted as 2.1/5 or 42 per cent rather than $\frac{2}{5}$ or 40 per cent. If it is the number eligible to retire that happens to be known exactly from another source, it should nevertheless be disregarded, and the proportion should be calculated from the sample numbers in all cases.

The conclusions stated in (a) and (b) above follow easily from a comparison of the coefficients of variation of p_2'/p_1', p_2/p_1', and p_2'/p_1 (or of their reciprocals), the dashes being used to indicate sample values. The expected value, variance, and coefficient of variation of p_2'/p_1' are given, respectively, by

$$E\left(\frac{p_2'}{p_1'}\right) = \frac{p_2}{p_1}$$

$$V\left(\frac{p_2'}{p_1'}\right) \doteq \frac{N-n}{Nn}\,\frac{p_2(p_1-p_2)}{p_1^3}$$

$$\left[C.V.\left(\frac{p_2'}{p_1'}\right)\right]^2 \doteq \frac{N-n}{Nn}\left(\frac{1}{p_2}-\frac{1}{p_1}\right)$$

where the approximations assume that n is sufficiently large.[1] If the reciprocal of p_2'/p_1' is expanded in a Taylor series in the neighborhood of the expected value of p_2'/p_1', it can be seen that, for sufficiently large n,

$$E\left(\frac{p_1'}{p_2'}\right) \doteq \frac{p_1}{p_2}$$

$$V\left(\frac{p_1'}{p_2'}\right) \doteq \frac{N-n}{Nn}\,\frac{p_1(p_1-p_2)}{p_2^3}$$

$$\left[C.V.\left(\frac{p_1'}{p_2'}\right)\right]^2 \doteq \left[C.V.\left(\frac{p_2'}{p_1'}\right)\right]^2 \doteq \frac{N-n}{Nn}\left(\frac{1}{p_2}-\frac{1}{p_1}\right).$$

Likewise, if the Taylor series expansion is applied to $1/p_1'$, it follows from the well known expressions for $E(p_1')$ and $V(p_1')$ that

[1] W. E. Deming, *Some Theory of Sampling*, pp. 452–54, quoting a proof by F. F. Stephan. Stephan's proof excludes the possibility of a zero value for the denominator proportion.

$$E\left(\frac{1}{p_1'}\right) \doteq \frac{1}{p_1}$$

$$V\left(\frac{1}{p_1'}\right) \doteq \frac{N-n}{Nn}\,\frac{1-p_1}{p_1^2}$$

$$\left[C.V.\left(\frac{1}{p_1'}\right)\right]^2 = \left[C.V.\left(\frac{p_2}{p_1'}\right)\right]^2 \doteq \frac{N-n}{Nn}\left(\frac{1}{p_1}-1\right).$$

For completeness, we have the familiar formula,

$$[C.V.(p_2')]^2 = \left[C.V.\left(\frac{p_2'}{p_1}\right)\right]^2 \doteq \frac{N-n}{Nn}\left(\frac{1}{p_2}-1\right).$$

The approximate formulas for the coefficients of variation are asymptotic expressions and so are the efficiency comparisons based on them.

2. A useful application of these results can often be made, even when the parameter under investigation is simply the proportion of cases in the universe possessing a specified attribute.[2] Thus, suppose the proportion (p_1) of A-items is desired, the proportion (p_2) of B-items is known, the B-items are a sub-class of the A-items, and $p_2 > p_1'/(2-p_1')$, or, alternatively, $p_1' < 2p_2/(1+p_2)$, (where p_1' is the sample value for p_1). Under these conditions, it is preferable to obtain the sample value (p_2') of p_2 from the same sample (assuming the B-items can conveniently be counted along with the A-items), and use $p_1'p_2/p_2'$ instead of p_1' as the estimate of p_1. The advantage of this procedure can be measured by the ratio

$$\frac{C.V.\left(\frac{p_1'p_2}{p_2'}\right)}{C.V.(p_1')} \doteq \left[\frac{p_1-p_2}{p_2(1-p_1)}\right]^{1/2}.$$

For example, if $p_1 = .2$ and $p_2 = .15$, $C.V.(p_1'p_2/p_2') \doteq .6\ C.V.\ (p_1')$.

Similarly, suppose the proportion (p_2) of B-items is desired, the proportion (p_1) of A-items is known, and the B-items are a sub-class of the A-items. It is preferable to obtain the sample value (p_1') of p_1 from the same sample (assuming the A-items can conveniently be counted along with the B-items), and use $p_2'p_1/p_1'$ instead of p_2' as the estimate of p_2. The advantage of this procedure can be measured by the ratio

$$\frac{C.V.\left(\frac{p_2'p_1}{p_1'}\right)}{C.V.(p_2')} \doteq \left[\frac{p_1-p_2}{p_1(1-p_2)}\right]^{1/2}.$$

For example, if $p_1 = .2$ and $p_2 = .15$, $C.V.(p_2'p_1/p_1') \doteq .5\ C.V.\ (p_2')$.

[2] See also W. E. Deming, *op. cit.*, pp. 165 ff.

THE CIRCULAR NORMAL DISTRIBUTION: THEORY AND TABLES

E. J. GUMBEL
Consultant, Stanford University

J. ARTHUR GREENWOOD
Manhattan Life Insurance Company

DAVID DURAND
National Bureau of Economic Research

0. PROBLEM

THE recognition of periodicities, their amplitude and length, is one of the main tasks of the application of statistics to long-range data and especially to economic phenomena. Although numerous theorems have been developed, the question of how to recognize a cycle remains open, since the complicated mathematical methods used in the periodogram and correlogram analysis may generate both real and factitious cycles and may conceal real ones.

Therefore, the study of cycles is a fertile field for nonscientific procedures. Since even a constant value in a certain interval can easily be reproduced by a Fourier series, cycles with superimposed epicycles and third and fourth cycles have been "found" in nearly every domain that may be represented by numbers. Different observers have found different cycles in the same series—not to mention the great cycles of history leading to forecasts, good advices and policies advocated by their prophets. This abuse has discredited the whole theory of cycles.

In contrast to these procedures, we examine here a specialized distribution function. Applications to recurrent phenomena of known period will be given in a second article. Since the distribution function has only two disposable parameters, and since we consider using it only in connection with such phenomena as seasonal or diurnal variations—where the length of period is definitely established from nonstatistical considerations—we feel that the chance of generating false cycles is small indeed.

The distribution function under discussion was tabulated because it possesses important theoretical properties of the linear normal distribution and because it has practical applications to such meteoric phenomena as the direction of the wind, which is an angular or "circular" variable. Wind directions are represented clockwise on the map from north to east, south, west, and back to north in the same way as an angle varies from zero to 2π. Each direction or angle has a

certain frequency and the sum of the frequencies is the number of observations, diminished by the number of calm conditions. The question then arose of extending the use of this function to time series with seasonal, or possibly diurnal, components. Assume n events happen during a year, or a span of several years. These events might be storms, or deaths, or the completion of automobiles. Each event happens at a certain date. If this date is considered as the chance variate α, the distribution of events over the year may be considered a circular distribution of an angle.

1. STATISTICS OF A CIRCULAR DISTRIBUTION

We first consider the purely empirical procedures which are used especially by climatologists in order to characterize circular distributions. These descriptive measures will prove basic for the construction of the theory and the estimation of the parameters.

Let n points be situated in some fashion on the circumference of a unit circle: then $\alpha_\nu (\nu = 1, 2, \cdots, n)$ are the angles of the radii drawn from the center. To characterize such circular observations their vector mean has been used: this is clearly identical with the centroid of the set of n points. To calculate it we introduce the rectangular coordinates $x_\nu = \cos \alpha_\nu$, $y_\nu = \sin \alpha_\nu$. Then

$$\bar{x} = \frac{1}{n} \sum_1^n x_\nu, \qquad \bar{y} = \frac{1}{n} \sum_1^n y_\nu \tag{1.1}$$

defines the vector mean.

For our purposes it is, however, more convenient to express this vector mean in polar coordinates. Let $\bar{a}$ and α_0 be a solution of the equations $\bar{x} = \bar{a} \cos \alpha_0$, $\bar{y} = \bar{a} \sin \alpha_0$. It is well known that this solution is unique unless $\bar{x} = \bar{y} = 0$, and is given by

$$\bar{a} = (\bar{x}^2 + \bar{y}^2)^{1/2} \tag{1.2}$$

$$\alpha_0 = \arctan \frac{\bar{y}}{\bar{x}} \tag{1.3}$$

where the quadrant in which α_0 lies must be determined by inspection of the signs of $\bar{x}$ and $\bar{y}$. Equation (1.2), which defines the vector strength $\bar{a}$ at the mean direction α_0, may be written from (1.1)

$$n\bar{a} = \sqrt{\left(\sum_1^n \sin \alpha_\nu \right)^2 + \left(\sum_1^n \cos \alpha_\nu \right)^2} \,. \tag{1.2'}$$

Finally the mean direction α_0 satisfies the equation

$$\sum_1^n \sin(\alpha_\nu - \alpha_0) = 0. \tag{1.3'}$$

The statistic α_0 is analogous to the average of a linear variate; the meaning of $\bar{a}$ will be explained later.

The observations α_ν may be given individually or, and this is mostly the case, in grouped form, i.e. as frequencies p_ν attributed to certain intervals, say months, of equal length.[1] In the individual case, we calculate the sine and the cosine for each observation, and obtain α_0 and $\bar{a}$ from (1.2') and (1.3). The grouped case permits a simplification. Assume that the frequencies for the 12 months are given, and assume furthermore months of thirty days each. Then one day corresponds to one degree. We attribute the frequencies to the midpoint of each month and assume, say, the middle of July as preliminary value for α_0. Then the observations can be arranged in the following scheme:

TABLE 1

SCHEME FOR A CIRCULAR DISTRIBUTION OVER THE YEAR

		North		
		360°		
	330° June	July	August	30°
	300° May		September	60°
West	270° April		October	90° East
	240° March		November	120°
	210° February		December	150°
		January		
		180°		
		South		

The mean direction is obtained from the following trigonometric rules:

$$\sum_{\nu=1}^n \cos\alpha_\nu = \text{July} - \text{Jan.} + 0.86603\,(\text{Aug.} - \text{Feb.} + \text{June} - \text{Dec.}) + 0.5\,(\text{Sept.} - \text{March} + \text{May} - \text{Nov.}) \tag{1.4}$$

$$\sum_{\nu=1}^n \sin\alpha_\nu = \text{Oct.} - \text{Apr.} + 0.86603\,(\text{Sept.} - \text{March} + \text{Nov.} - \text{May}) + 0.5\,(\text{Aug.} - \text{Feb.} + \text{Dec.} - \text{June}). \tag{1.5}$$

[1] Whether corrections for length of month are desirable in practical calculations will depend on the degree of refinement desired, and on the amplitude of the seasonal movements. When the seasonals are of the order of 10 per cent or less, variations in length of month will produce noticeable irregularities. February, after all, has 10 per cent fewer days than January. The details of the corrections will be given in the second article.

Of course this procedure can also be applied to any other starting point after the appropriate rotations.

A circular distribution may be graphically represented by a linear histogram. Numerous arrangements are possible, including:

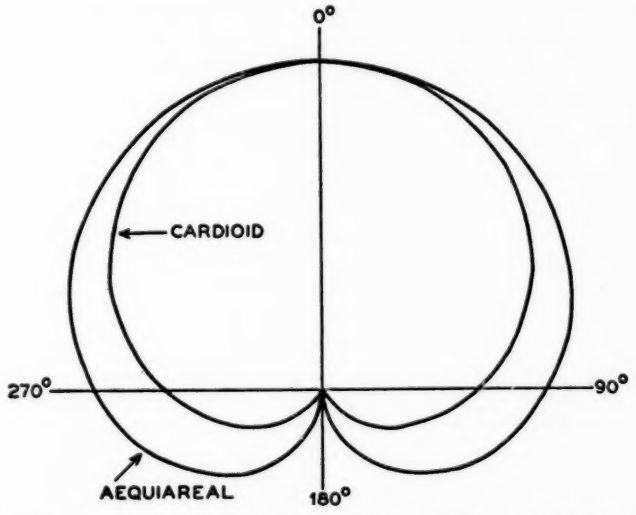
(1) Listing the months in their normal order from January on the left to December on the right;
(2) Placing the modal month in the middle with one-half of the sixth following month at either end;
(3) Locating the calculated mean (or mode) in the middle.

A circular distribution may also be plotted on polar coordinate paper. Here the usual procedure is to trace the frequencies of the various months as radii vectores at the appropriate angles; the result is a polar wedge diagram corresponding to the linear histogram. By this procedure, however, the area of each wedge varies as the square of the frequency, which means that if two different distributions of, say, 100 observations are plotted, their areas will in general differ. To conserve areas, Leighly [8] proposes to trace the square roots of the frequencies, and this procedure, henceforth called aequiareal, is used in the following. An obvious alternative is to use aequiareal polar paper, which may be constructed by graduating the radii according to a square root scale.

Two limiting cases of circular distributions are immediately evident. The first is the uniform distribution, where any angle is as likely to occur as any other. Then the density of probability $f(\alpha) = \frac{1}{2}\pi$ (α measured in radians) is obviously independent of α. The cumulative probability function—defined in analogy to the linear procedure by $dF/d\alpha = f(\alpha)$, with the boundary conditions $F(0) = 0$ and $F(2\pi) = 1$—is simply $F(\alpha) = \alpha/2\pi$. Pólya used an n-dimensional analogue of this assumption when he investigated whether the stars are distributed at random over the celestial sphere—also when he calculated the mean distance between two points for any number of dimensions [14, 16].

The other limiting case occurs when all the angles are the same. This, of course, has no practical or theoretical interest, but the asymptotic approach to this case is worth discussion. As an example, a set of angles might be normally distributed with mean 0 and standard deviation 1°. Then no sensible fraction of the total frequency would be found in most of the circle—for example, the arc from 5° to 355° would be virtually void. Asymptotic distributions of this type do not differ sensibly from the linear normal. They have been used for the analysis of the Brownian movement and the paths of the beta rays [6], and they are the basis for Schelling's work on the most frequent particle paths in two and three dimensions [18].

The problem of circular distributions becomes significant when a sensible portion of the observations may occur over all or nearly all of the circumference. As in the linear case an infinity of circular distributions exists. Among them a few, possessing common desirable properties, have been studied. The distributions studied are all symmetrical, although many observed phenomena show asymmetry. Moreover, these distributions possess one mode—i.e., one most probable direction. From both properties—symmetry and unimodality—it follows that nowhere is the probability less than its value at the antipode of the mode—



GRAPH 1. The cardioid and the aequiareal trace of the cardioid distribution.

henceforth called the antimode. Nothing is lost in generality if we put the mode at zero, north on the map, and the antimode at π, south.

A simple example of a circular distribution is the sine function modified in such a way that the density of probability is non-negative and the area under the distribution is unity. This leads to the symmetrical distribution

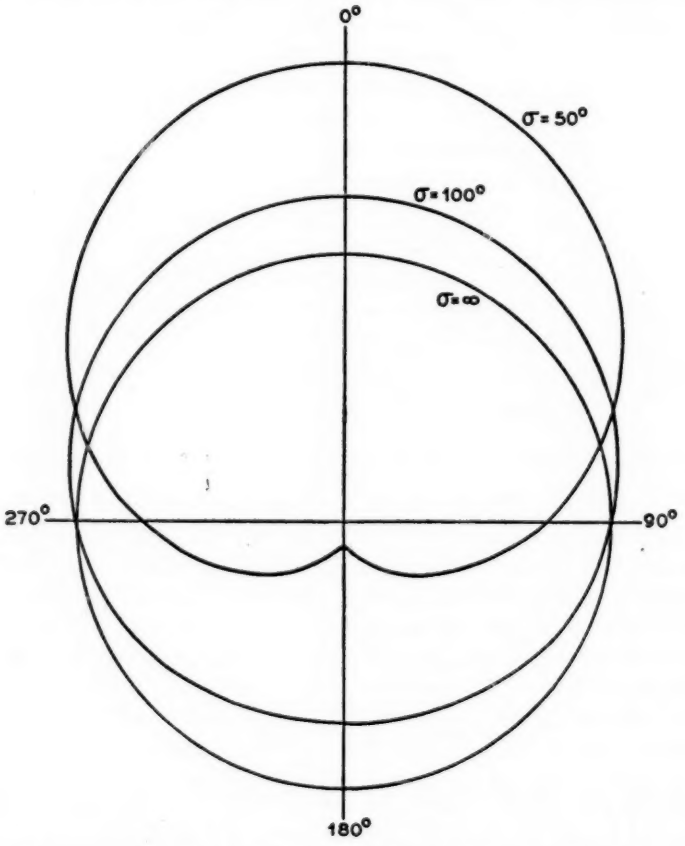
$$f(\alpha) = \tfrac{1}{2}(1+\cos\alpha).$$

The density of probability is $1/\pi$ at the mode and zero at the antimode. When $f(\alpha)$ is traced in polar coordinates, the resulting curve is the cardioid. Graph 1 compares the cardioid trace of $f(\alpha)$ with the aequiareal trace.

Other circular distributions can be generated by wrapping an unlimited linear distribution around a circumference. This problem has been studied by Lévy, Marcinkiewicz, and Wintner [9, 10, 20], who have thus treated the normal and the Cauchy distributions. The normal wrapped-up distribution with mode zero may be written

$$(1.6)\quad \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=-\infty}^{+\infty} e^{-(\alpha+2k\pi)^2/2\sigma^2} = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{t=1}^{\infty} q^{t^2} \cos t\alpha \quad (0 \leqq \alpha \leqq 2\pi)$$

where $\ln q = -\sigma^2/2$. Graph 2 shows it traced on aequiareal polar scale for $\sigma = 50°$ and $100°$, corresponding respectively to $q = .68333383$ and $.21608421$. When $\sigma > 77°$ approximately, this angular distribution may be computed from tables of elliptic theta functions. Theoretical advan-



GRAPH 2. Wintner's circular distribution in aequiareal scale.

tages recommend this distribution; it presents, however, serious difficulties, especially if the method of maximum likelihood is to be used to estimate its parameters.

2. THE CIRCULAR NORMAL DISTRIBUTION

Gauss has shown that the normal distribution can be derived by what is now called the principle of maximum likelihood and a single assumption that the mean is the most probable value. In 1918 Mises [11] applied the method of Gauss to a circular variate and derived the distribution that is the subject of this paper. His procedure is essentially as follows: Mises asks for such a distribution $f(\xi)$, where ξ is an "error of observation" $\xi_\nu = \alpha_\nu - \alpha$, that the ratio of the *a posteriori* to the *a priori* probability (the likelihood function) of a "true value" α upon n observations $\alpha_1, \cdots, \alpha_n$ is to be a maximum for that α given by equation (1.3′). The likelihood function is

$$\prod_{\nu=1}^{n} f(\xi_\nu).$$

The postulate means that

$$\sum_{\nu=1}^{n} \frac{f'(\alpha_\nu - \alpha_0)}{f(\alpha_\nu - \alpha_0)} = 0$$

should hold together with equation (1.3′). Since these two sums are equal for arbitrary values of the α_ν, the equality must hold term by term. This functional equation has the solution

$$f(\xi_\nu) = Ce^{k \cos \xi_\nu}$$

where the two positive parameters C and k are linked by the condition

$$\int_0^{2\pi} f(\xi)d\xi = 1.$$

Consequently

$$C = 1 \Big/ \int_0^{2\pi} e^{k \cos \xi} d\xi$$

$$= 1/2\pi I_0(k)$$

where $I_0(k)$ is the Bessel function of the first kind of pure imaginary argument. Therefore

$$f(\alpha) = \frac{e^{k \cos (\alpha - \alpha_0)}}{2\pi I_0(k)}. \tag{2.1}$$

This expression is henceforth called the *circular normal distribution* since it was derived in a way which is strictly analogous to the classical derivation of the linear normal distribution. Pólya [15, pp. 129–132] has shown that the distribution (2.1) is the only one that shares certain important properties of the linear normal.

The density function has a mode at α_0 and an antimode at $\alpha_0+\pi$. The quotient of the densities at the mode and at the antimode is

$$f(\alpha_0)/f(\alpha_0 + \pi) = e^{2k}. \tag{2.2}$$

The larger the part of the distribution situated in the neighborhood of the mode, the larger is k. Therefore, this parameter k is a measure of concentration. For $k=0$ the distribution (2.1) degenerates into the uniform circular distribution. When k is significantly different from zero, there is evidence of the existence of a cycle.

To get further insight into the nature of the parameter k consider the points of inflection for the distribution traced on a linear scale. Let the mode α_0 be zero, then the points of inflection α_1 and α_2 are the solutions of

$$f''(\alpha) = 0$$

whence from (2.1)

$$k \sin^2 \alpha = \cos \alpha.$$

The larger the parameter k the nearer are the points of inflection to the mode. The preceding equation has the real solutions

$$\cos \alpha_1 = -\frac{1}{2k} + \sqrt{1 + \frac{1}{4k^2}}; \alpha_2 = -\alpha_1. \tag{2.3}$$

For large values of k, the points of inflection are approximately

$$\alpha_{1,2} = \pm 1/\sqrt{k} \tag{2.3'}$$

while for the linear normal distribution, the points of inflection are

$$\alpha_{1,2} = \pm \sigma.$$

If k is large, its reciprocal $1/k$ influences the circular normal distribution in the same way as σ^2 influences the linear normal one.

This analogy leads to the second degenerate case. Let k be large; then the distribution is highly concentrated about α_0. Introduce a reduced variate β by writing

$$\alpha - \alpha_0 = \beta/\sqrt{k}.$$

The distribution of the variate β is

$$f(\beta) = \text{const } e^{k \cos (\beta/\sqrt{k})}.$$

Since the argument of the cosine is small, we may neglect higher powers of β in the exponential and obtain a normal distribution

$$f(\beta) = \frac{1}{\sqrt{2\pi}} e^{-\beta^2/2} \tag{2.4}$$

with mean zero and unit standard deviation. In fact, for large k the reduced distribution $f(\beta)$ converges to this linear normal distribution.

For the estimation of the parameters consider the likelihood function

$$L = (2\pi)^{-n} I_0^{-n}(k) \exp\left[k \sum_1^n \cos(\alpha_\nu - \alpha_0)\right].$$

Equating $\partial/\partial\alpha_0 \log L$ to zero we obtain equation (1.3′). Equating $\partial/\partial k \log L$ to zero we obtain

$$\frac{I_0'(k)}{I_0(k)} = \frac{1}{n} \sum_1^n \cos(\alpha_\nu - \alpha_0).$$

The expression on the right is easily transformed with the help of (1.3′) into

$$\frac{1}{n} \sqrt{\left(\sum_1^n \cos \alpha_\nu\right)^2 + \left(\sum_1^n \sin \alpha_\nu\right)^2}$$

which by (1.2′) equals the vector strength $\bar{a}$. Therefore the parameter k is the solution of

$$I_1(k) - \bar{a} I_0(k) = 0 \tag{2.5}$$

where

$$I_1(k) = I_0'(k).$$

Equation (2.5) was solved for k by interpolating in a table of Bessel functions [16a]. Table 2 gives k as a function of the observed vector strength $\bar{a}$. Tables for testing the significance of observed $\bar{a}$ are under preparation.

K. Arnold [1] has generalized the distribution (2.1) to the bivariate case, considering the center of gravity of points on a sphere. He has also generalized equation (1.6) to the bivariate case by using Fourier's

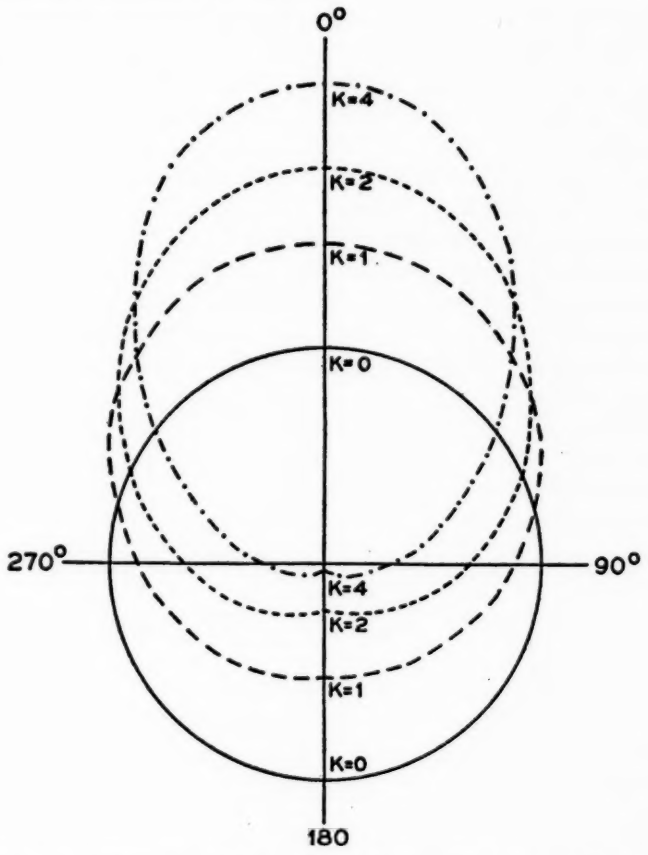
TABLE 2

PARAMETER k AS A FUNCTION OF VECTOR STRENGTH $\bar{a}$

$$\bar{a} = \frac{I_1(k)}{I_0(k)}$$

a	k	δ^2	$\bar{a}$	k	δ^2
.00	0.00000	+	.45	1.01022	+52
.01	.02000	1	.46	.03889	54
.02	.04001	1	.47	.06810	57
.03	.06003	1	.48	.09788	62
.04	.08006	4	.49	.12828	64
.05	.10013	2	.50	.15932	69
.06	.12022	3	.51	.19105	72
.07	.14034	5	.52	.22350	77
.08	.16051	5	.53	.25672	83
.09	.18073	6	.54	.29077	88
.10	0.20101	5	.55	1.32570	93
.11	.22134	8	.56	.36156	100
.12	.24175	7	.57	.39842	107
.13	.26223	8	.58	.43635	115
.14	.28279	9	.59	.47543	123
.15	.30344	10	.60	.51574	133
.16	.32419	9	.61	.55738	142
.17	.34503	12	.62	.60044	156
.18	.36599	12	.63	.64506	166
.19	.38707	13	.64	.69134	183
.20	0.40828	13	.65	1.73945	197
.21	.42962	14	.66	.78953	216
.22	.45110	15	.67	.84177	236
.23	.47273	17	.68	.89637	260
.24	.49453	16	.69	.95357	285*
.25	.51649	18	.70	2.01363	315*
.26	.53863	20	.71	.07685	351*
.27	.56097	19	.72	.14359	391*
.28	.58350	22	.73	.21425	437*
.29	.60625	22	.74	.28930	493*
.30	0.62922	23	.75	2.36930	558*
.31	.65242	25	.76	.45490	633*
.32	.67587	26	.77	.54686	728*
.33	.69958	27	.78	.64613	837*
.34	.72356	29	.79	.75382	972*
.35	.74783	31	.80	.87129	1136*
.36	.77241	31	.81	3.00020	1341*
.37	.79730	34	.82	.14262	1598*
.38	.82253	36	.83	.30114	1918*
.39	.84812	37	.84	.47901	2331*
.40	0.87408	39	.85	3.68041	2860*
.41	.90043	42	.86	.91072	3557*
.42	.92720	43	.87	4.17703	4480*
.43	.95440	47			
.44	.98207	48			
.45	1.01022	52			

* δ has been modified from .69 to .87.



GRAPH 3. The circular normal distribution in aequiareal scale for $k=0, 1, 2, 4$.

equation of heat flow, which yields the normal linear equation as a special case.

3. CALCULATION OF TABLES

The density function (2.1) is relatively simple to calculate from tables of the Bessel function, the cosine, and the exponential function. The density function is presented to three decimals in the form

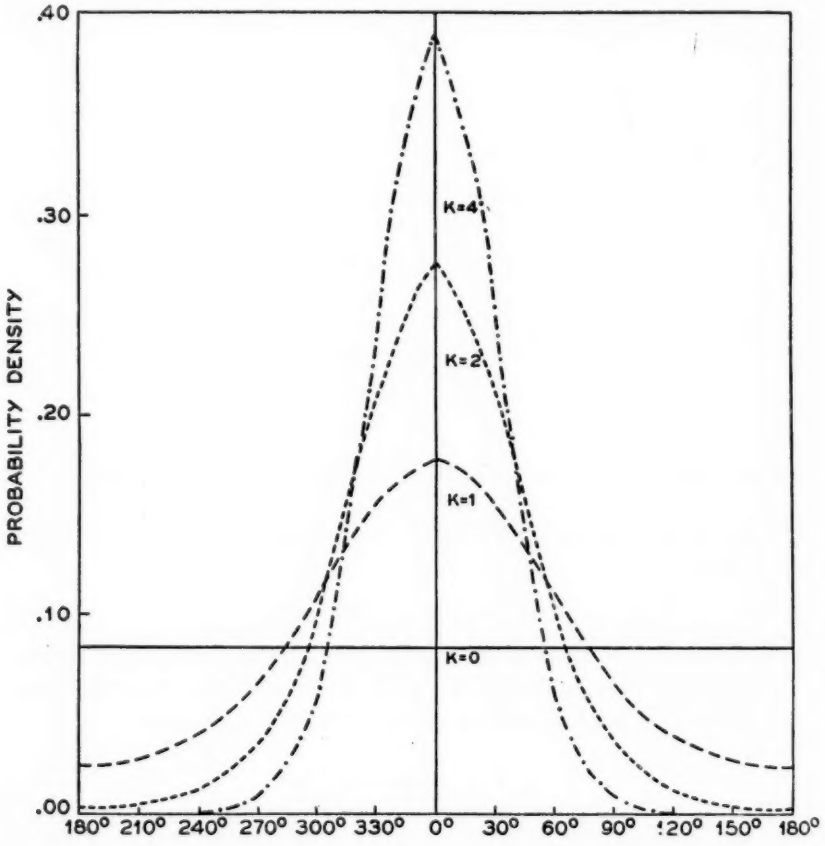
$$\psi(\alpha) = \sqrt{e^{k\cos\alpha}/I_0(k)} \tag{3.1}$$

which is adequate for tracing aequiareal frequency ovals on polar paper (Table 3). This table was calculated to five decimals and checked

by differencing before cutting down to three decimals. It was further checked by multiplying together the values corresponding to supplementary angles. For tracing, the values given in Table 3 must be multiplied by $\sqrt{n/m}$, where m stands for the number of wedges—in our case $m=12$.

Graph 3 traced in aequiareal polar scale shows the influence of the parameter k. For increasing values of k the area from 90 to 270 degrees shrinks to an insignificant part of the whole area. For the sake of comparison the distributions are also drawn in linear scale in Graph 4.

The areas of the circular normal distribution function, Table 4, posed the main problem of calculation. Originally it was desired to compute this table to 7 decimals for $k=0.1$ (0.1) 4.0 and $\alpha=2\frac{1}{2}°$ ($2\frac{1}{2}°$)



GRAPH 4. The circular normal distribution in linear scale for $k=0, 1, 2, 4$.

180°. This was actually done, and the present table is an abridgment. For calculation three methods were considered: quadratures, expansion in a series of incomplete beta functions, and expansion in a series of sines. The sine expansion was finally accepted—partly because it was suitable for use with an I.B.M. 602 calculating punch available to the authors, and partly because adequate tables of the Bessel functions $I_n(k)$ were obtainable.

The method of quadratures was suggested in 1950 by Dr. H. R. J. Grosch, then of the Watson Scientific Computing Laboratory, Columbia University. The method has the advantage that the complete integral, $\int_0^\pi e^{k\cos x}\,dx$, is obtained without reference to its definition as a Bessel function, and the Bessel function is then used as a check. This method would be suitable for longhand calculation, and it seemed ideally suited to calculation on the card programmed calculator at the Watson Laboratory—a possibility briefly considered. The method was not, however, well adapted to calculation on the 602 calculating punch.

The method of incomplete beta functions derives from Mises [11, p. 499], who gives the formula

$$\int_0^\alpha e^{k\cos x}dx = \sum_{n=0}^{\infty} \frac{k^n}{n!}\int_0^\alpha \cos^n x\,dx \tag{3.2}$$

and indicates that it was used for computation by S. Weich. Arnold has likewise stated, in conversation with the authors, that he used a formula equivalent hereto in computing his 6 decimal table [1]. The integral on the right side of (3.2) is recognized as the incomplete beta function

$$\frac{1}{2}\int_0^v t^{-1/2}(1-t)^{(n-1)/2}\,dt, \quad \text{where} \quad v = \sin^2 x.$$

To evaluate it there are available, besides expansion of the integrand as a polynomial in $\cos mx$:

(1) The tables of Legendre [7, p. 178] available only for $n=2, 4, 6$.

(2) The table of Soper [12, p. 179], for $x=\pi/2$. This was used in preliminary calculations.

(3) The table of Wishart [12, p. 239]. This was investigated but its auxiliary argument, which we shall call ξ, defined by

$$\xi = 2n^{1/2}\tan\frac{x}{2},$$

renders it unsuitable for extended calculations for fixed x; moreover it fails for $n<9$.

(4) The extensive beta-function table of Pearson [13]. This requires the auxiliary variable $z=\cos^2 x$, rendering interpolation everywhere necessary and requiring, for moderately small x, some such artifice as that of Pearson and Stoessiger [12, p. 176].

The method of sine series rests on the expansion

$$e^{k\cos x} = I_0(k) + 2\sum_{n=1}^{\infty} I_n(k)\cos nx \tag{3.3}$$

which was essentially given by Frullani [4, p. 502]. Integration term by term gives

$$\frac{\int_0^{\alpha} e^{k\cos x}dx}{\int_0^{\pi} e^{k\cos x}dx} = \frac{\alpha}{\pi} + \frac{2}{\pi I_0(k)}\sum_{n=1}^{\infty} n^{-1}I_n(k)\sin n\alpha. \tag{3.4}$$

When the tabulation of the distribution by the above formula was first considered, the then most readily available tables of $I_n(x)$ were those of A. Lodge [5, p. 309]. The interpolation of these for intervals of 0.1 and the extrapolation by recurrence to I_{14}, which preliminary calculation showed would be necessary for 9 decimal accuracy throughout, were not works we were prepared to undertake. Then Dr. J. C. P. Miller generously provided us with proof sheets of the Royal Society's forthcoming volume of Bessel functions of integral order, proceeding at intervals of 0.1 and going up to I_{18}.

The connection between the two series expansions should be briefly noted. We have

$$I_0(k) + 2\sum_{n=1}^{\infty} I_n(k)\cos nx = \sum_{n=0}^{\infty}\epsilon_n\cos nx\sum_{t=0}^{\infty}\frac{k^{n+2t}}{2^{n+2t}t!(n+t)!} \tag{3.5}$$

$$(\epsilon_0 = 1,\ \epsilon_n = 2,\ n > 0)$$

$$\sum_{n=0}^{\infty}\frac{k^n}{n!}\cos^n x = \sum_{n=0}^{\infty}\frac{k^n}{2^n}\sum_{t=0}^{(n/2)}\frac{\epsilon_{n-2t}\cos(m-2t)x}{t!(n-t)!}. \tag{3.6}$$

Hence, if we take the beta function series and neglect terms beyond that in k^n, it is equivalent to replacing each Bessel function by its ascending power series and truncating these at the terms in k^n. The double series converge absolutely, and so we have incidentally proved that the coefficients of the cosine series are the I_n; this was not the method used by Frullani, who, writing before the modified Bessel functions had

been defined, obtained the differential equation satisfied by the coefficient of cos nx by integrating by parts and thence derived the power series.

The use of punch card equipment to calculate series such as (3.4) is now fairly widespread. According to Comrie [2], "The first scientific application of the Hollerith—in 1928—was to the summation of harmonic terms, or, in other words, Fourier synthesis." Comrie's description of this job [3] is still worth reading; the reader should note carefully the difference between English and American machines [2, pp. 156–157] and remember that the 80-column card is now nearly universal.

In our calculations, values of $2/\pi I_0(k)$ were determined by hand calculator to 10 decimals and punched into cards for $k = [0.1\ (0.1)\ 4.2]$. I_n and n were punched; then $2I_n(k)/\pi I_0(k)$ and $2I_n(k)/n\pi I_0(k)$ were formed successively on the 602 calculating punch. The angles nx, $1 \leqq n \leqq 14$, $x = [0°\ (2.5°)\ 180°]$, were reduced by hand to the first quadrant and the sines were punched to 9 decimals from the National Bureau of Standards table [19].

The term α/π, being a multiple of 1/72, was calculated on an I.B.M. 405 tabulator by progressive summation. The sines were multiplied by the Bessel function quotients on the 602. Although the summations could have been made on this machine also, they were made on a tabulator (both 405 and 402 were used) in order to have a printed record of the terms; the process differs from that in [2] principally in the fact that the tabulators used in our calculations are equipped to subtract from the card. A 513 summary punch was connected to punch a 9 decimal table of the function (with errors up to 7 units in the last place); this table was printed and examined for gross errors. It was then checked by forming first to sixth differences, inclusive, in both the k- and α-directions, on the 602. From the 9 decimal second and fourth differences, in each direction, modified second differences were computed by the usual formula

$$\delta_m^2 = \delta^2 - .18393\delta^4.$$

The function and the modified differences in the two directions were then rounded to 7 decimals by adding 5 to the eighth decimal. The modification and the rounding were checked by summing the cards in groups on the tabulator, and the table was then printed.

The present abridgment, Table 4, was prepared by taking cards corresponding to increments of 0.2 in k and 5 degrees in α and differencing them five times in each direction; the fifth differences were used only to

TABLE 3

RADII OF PROBABILITY DENSITY FOR AEQUIAREAL TRACING OF THE CIRCULAR NORMAL DISTRIBUTION

$$\psi(\alpha) = \sqrt{\frac{e^{k\cos\alpha}}{I_0(k)}}$$

K	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°	100°	110°	120°	130°	140°	150°	160°	170°	180°	K
0.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.0
0.1	1.050	1.049	1.047	1.043	1.038	1.031	1.024	1.016	1.007	0.999	0.990	0.981	0.974	0.957	0.961	0.956	0.953	0.951	0.950	0.1
0.2	1.100	1.098	1.093	1.085	1.074	1.061	1.046	1.030	1.012	0.995	0.978	0.962	0.947	0.933	0.922	0.912	0.906	0.902	0.900	0.2
0.3	1.149	1.146	1.139	1.126	1.109	1.089	1.066	1.041	1.015	0.989	0.963	0.939	0.917	0.898	0.882	0.868	0.859	0.853	0.851	0.3
0.4	1.197	1.194	1.183	1.166	1.143	1.115	1.084	1.050	1.015	0.980	0.947	0.916	0.887	0.862	0.841	0.824	0.812	0.805	0.803	0.4
0.5	1.245	1.240	1.226	1.204	1.174	1.139	1.099	1.056	1.013	0.970	0.929	0.890	0.856	0.826	0.801	0.781	0.767	0.758	0.755	0.5
0.6	1.292	1.286	1.269	1.241	1.204	1.160	1.112	1.060	1.008	0.957	0.908	0.864	0.824	0.789	0.760	0.738	0.722	0.712	0.709	0.6
0.7	1.337	1.330	1.309	1.276	1.232	1.180	1.122	1.062	1.001	0.942	0.887	0.836	0.791	0.752	0.721	0.696	0.678	0.668	0.664	0.7
0.8	1.381	1.373	1.348	1.309	1.258	1.197	1.131	1.062	0.992	0.926	0.864	0.807	0.758	0.716	0.682	0.655	0.624	0.624	0.621	0.8
0.9	1.424	1.414	1.386	1.341	1.282	1.213	1.137	1.059	0.982	0.908	0.840	0.778	0.725	0.680	0.643	0.615	0.595	0.583	0.579	0.9
1.0	1.465	1.454	1.422	1.370	1.304	1.226	1.141	1.054	0.969	0.889	0.815	0.749	0.692	0.644	0.606	0.576	0.556	0.543	0.539	1.0
1.1	1.505	1.493	1.456	1.398	1.323	1.237	1.143	1.048	0.955	0.868	0.789	0.719	0.660	0.610	0.570	0.539	0.518	0.505	0.501	1.1
1.2	1.543	1.529	1.489	1.424	1.341	1.246	1.143	1.040	0.940	0.847	0.763	0.690	0.628	0.576	0.535	0.504	0.482	0.469	0.465	1.2
1.3	1.580	1.565	1.520	1.449	1.357	1.253	1.142	1.030	0.924	0.825	0.737	0.661	0.596	0.543	0.501	0.470	0.448	0.435	0.431	1.3
1.4	1.616	1.599	1.549	1.471	1.372	1.258	1.139	1.019	0.906	0.802	0.711	0.632	0.565	0.512	0.469	0.438	0.416	0.403	0.398	1.4
1.5	1.650	1.631	1.577	1.492	1.384	1.262	1.134	1.007	0.888	0.779	0.684	0.603	0.536	0.481	0.439	0.407	0.385	0.372	0.368	1.5
1.6	1.682	1.662	1.603	1.511	1.395	1.264	1.128	0.994	0.869	0.756	0.658	0.575	0.507	0.452	0.410	0.378	0.356	0.344	0.340	1.6
1.7	1.714	1.692	1.628	1.529	1.405	1.265	1.120	0.980	0.849	0.732	0.632	0.548	0.479	0.424	0.382	0.351	0.330	0.317	0.313	1.7
1.8	1.744	1.720	1.652	1.546	1.413	1.264	1.112	0.965	0.829	0.709	0.606	0.521	0.452	0.398	0.356	0.325	0.304	0.292	0.288	1.8
1.9	1.773	1.747	1.674	1.561	1.419	1.263	1.102	0.949	0.809	0.686	0.581	0.495	0.426	0.372	0.331	0.301	0.281	0.269	0.265	1.9
2.0	1.800	1.773	1.695	1.575	1.425	1.260	1.092	0.932	0.788	0.662	0.557	0.470	0.402	0.348	0.308	0.279	0.259	0.247	0.244	2.0
2.1	1.827	1.798	1.715	1.587	1.429	1.256	1.081	0.916	0.767	0.639	0.533	0.446	0.378	0.326	0.286	0.258	0.238	0.227	0.224	2.1
2.2	1.853	1.822	1.734	1.599	1.432	1.251	1.069	0.898	0.747	0.617	0.509	0.423	0.356	0.304	0.266	0.238	0.219	0.209	0.205	2.2
2.3	1.877	1.845	1.752	1.609	1.435	1.245	1.056	0.881	0.726	0.594	0.487	0.401	0.335	0.284	0.246	0.220	0.202	0.192	0.188	2.3
2.4	1.901	1.867	1.769	1.619	1.436	1.238	1.043	0.863	0.705	0.573	0.465	0.380	0.314	0.265	0.228	0.203	0.185	0.176	0.172	2.4
2.5	1.924	1.888	1.785	1.628	1.436	1.231	1.030	0.845	0.685	0.551	0.444	0.360	0.295	0.247	0.212	0.187	0.170	0.161	0.158	2.5
2.6	1.947	1.908	1.800	1.635	1.436	1.223	1.016	0.828	0.665	0.531	0.423	0.340	0.277	0.230	0.196	0.172	0.156	0.147	0.145	2.6
2.7	1.968	1.928	1.814	1.642	1.435	1.215	1.002	0.810	0.645	0.510	0.404	0.322	0.260	0.214	0.181	0.158	0.143	0.135	0.132	2.7
2.8	1.989	1.947	1.828	1.649	1.433	1.206	0.968	0.792	0.625	0.490	0.385	0.304	0.244	0.199	0.168	0.146	0.132	0.124	0.121	2.8
2.9	2.009	1.965	1.841	1.654	1.431	1.197	0.973	0.774	0.606	0.471	0.366	0.287	0.228	0.186	0.155	0.134	0.121	0.113	0.111	2.9
3.0	2.029	1.983	1.853	1.659	1.428	1.187	0.958	0.756	0.587	0.453	0.349	0.271	0.214	0.173	0.143	0.123	0.111	0.103	0.101	3.0
3.1	2.048	2.000	1.865	1.664	1.425	1.177	0.943	0.738	0.569	0.435	0.332	0.256	0 200	0.160	0.133	0.114	0.101	0.094	0.092	3.1
3.2	2.066	2.016	1.876	1.667	1.421	1.167	0.928	0.721	0.551	0.417	0.316	0.241	0.187	0.149	0.122	0.104	0.093	0.088	0.084	3.2
3.3	2.084	2.032	1.887	1.671	1.417	1.156	0.913	0.704	0.533	0.400	0.301	0.228	0.175	0.139	0.113	0.096	0.085	0.079	0.077	3.3
3.4	2.102	2.048	1.897	1.673	1.412	1.145	0.898	0.687	0.516	0.384	0.286	0.215	0.164	0.129	0.104	0.088	0.078	0.072	0.070	3.4
3.5	2.119	2.063	1.906	1.676	1.407	1.134	0.883	0.670	0.499	0.368	0.272	0.202	0.153	0.120	0.096	0.081	0.071	0.066	0.064	3.5
3.6	2.135	2.078	1.916	1.678	1.401	1.123	0.868	0.653	0.482	0.353	0.258	0.191	0.144	0.111	0.089	0.074	0.065	0.060	0.058	3.6
3.7	2.151	2.092	1.924	1.679	1.396	1.111	0.853	0.637	0.466	0.338	0.245	0.180	0.134	0.103	0.082	0.068	0.059	0.055	0.053	3.7
3.8	2.167	2.106	1.933	1.680	1.390	1.099	0.838	0.621	0.451	0.324	0.233	0.169	0.125	0.096	0.076	0.063	0.054	0.050	0.048	3.8
3.9	2.183	2.119	1.941	1.681	1.383	1.088	0.823	0.605	0.436	0.311	0.221	0.159	0.117	0.089	0.070	0.057	0.050	0.046	0.044	3.9
4.0	2.198	2.132	1.948	1.681	1.377	1.076	0.809	0.590	0.421	0.297	0.210	0.150	0.109	0.082	0.064	0.053	0.045	0.042	0.040	4.0

TABLE 4

AREAS OF THE CIRCULAR NORMAL DISTRIBUTION

$$\Phi(\alpha) = \int_{-\alpha}^{\alpha} f(x)\,dx$$

α	$k=0.0$			$k=0.2$			$k=0.4$			$k=0.6$			$k=0.8$			$k1.0$		
	Φ_α	δ_{α^2}	δ_{k^2}	Φ_α	δ_{α^2}	δ_{k^2}	Φ_α	δ_{α^2}	δ_{k^2}	Φ_α	δ_{α^2}	δ_{k^2}	Φ_α	δ_{α^2}	δ_{k^2}	Φ_α	δ_{α^2}	δ_{k^2}
	.00000	0		.00000	0		.00000	0		.00000	0		.00000	0		.00000	0	
5	.02778	0	56	.03358	− 5	− 43	.03981	− 12	28	.04631	− 21	12	.05294	− 32	− 1	.05956	− 45	− 12
10	.05556	0	110	.06711	−10	83	.07950	− 24	53	.09242	− 42	23	.10556	− 64	− 4	.11868	− 89	− 26
15	.08333	0	160	.10054	−15	120	.11895	− 36	75	.13810	− 62	29	.15755	− 93	− 11	.17690	−131	− 43
20	.11111	0	206	.13382	−20	152	.15804	− 46	91	.18317	− 80	30	.20861	−121	− 24	.23383	−168	− 66
25	.13889	0	246	.16691	−24	177	.19668	− 57	100	.22744	− 97	24	.25846	−145	− 43	.28908	−201	− 94
30	.16667	0	278	.19974	−29	194	.23474	− 66	100	.27075	−112	10	.30687	−167	− 69	.34234	−228	−129
35	.19444	0	301	.23230	−32	202	.27215	74	92	.31293	−125	− 13	.35362	−184	−102	.39333	−250	−169
40	.22222	0	316	.26453	−36	200	.30882	− 81	75	.35387	−136	− 43	.39853	−198	−142	.44183	−265	−215
45	.25000	0	321	.29640	−39	189	.34467	− 88	50	.39346	−144	− 80	.44148	−207	−187	.48768	−275	−265
50	.27778	0	316	.32788	−42	169	.37965	− 92	15	.43161	−150	−125	.48236	−213	−238	.53080	−279	−317
55	.30556	0	301	.35894	−44	140	.41371	− 96	− 26	.46826	−154	−174	.52111	−216	−291	.57113	−279	−370
60	.33333	0	278	.38956	−46	102	.44681	− 99	− 73	.50337	−156	−227	.55771	−215	−346	.60868	−274	−422
65	.36111	0	246	.41972	−47	59	.47893	−100	−125	.53692	−156	−283	.59217	−212	−400	.64350	−265	−472
70	.38889	0	206	.44941	48	9	.51004	−101	−180	.56893	−154	−338	.62451	−206	−451	.67567	−254	−517
75	.41667	0	160	.47862	−49	− 44	.54015	−100	−236	.59939	−150	−392	.65480	−198	−499	.70531	−240	−556
80	.44444	0	110	.50734	−49	− 99	.56927	− 98	−291	.62836	−146	−442	.68312	−188	−542	.73256	−224	−589
85	.47222	0	56	.53557	−49	−155	.59740	− 96	−344	.65587	−140	−488	.70956	−178	−578	.75757	−208	−615
90	.50000	0	0	.56331	−48	−210	.62457	− 93	−393	.68199	−133	−528	.73422	−166	−607	.78049	−191	−632
95	.52778	0	− 56	.59057	−47	−262	.65081	− 90	−436	.70677	−126	−560	.75722	−154	−627	.80150	−175	−642
100	.55556	0	−110	.61737	−46	−309	.67615	− 86	−473	.73029	−118	−585	.77868	−142	−640	.82077	−158	−643
105	.58333	0	−160	.64371	−44	−351	.70064	− 81	−503	.75263	−110	−602	.79871	−131	−644	.83844	−143	−637
110	.61111	0	−206	.66960	−42	−385	.72431	− 76	−524	.77387	−102	−609	.81744	−119	−640	.85469	−128	−623
115	.63889	0	−246	.69508	−40	−412	.74722	− 71	−537	.79410	− 94	−608	.83498	−107	−628	.86966	−114	−603
120	.66667	0	−278	.72016	−38	−430	.76942	− 66	−540	.81338	− 85	−599	.85144	− 96	−608	.88349	−100	−577
125	.69444	0	−301	.74486	−35	−438	.79096	− 61	−534	.83181	− 77	−581	.86694	− 86	−581	.89632	− 88	−545
130	.72222	0	−316	.76921	−32	−437	.81189	− 55	−519	.84947	− 69	−555	.88157	− 76	−548	.90826	− 77	−509
135	.75000	0	−321	.79324	−29	−427	.83227	− 50	−495	.86643	− 62	−521	.89545	− 67	−509	.91943	− 67	−468
140	.77778	0	−316	.81697	−26	−407	.85215	− 44	−463	.88278	− 54	−481	.90865	− 58	−465	.92993	− 57	−424
145	.80556	0	−301	.84044	−23	−378	.87159	− 39	−423	.89858	− 47	−434	.92128	− 49	−416	.93986	− 48	−376
150	.83333	0	−278	.86367	−20	−341	.89065	− 33	−376	.91392	− 40	−382	.93341	− 42	−363	.94930	− 40	−327
155	.86111	0	−246	.88670	−17	−296	.90938	− 27	−323	.92886	− 33	−326	.94513	− 34	−308	.95834	− 33	−275
160	.88889	0	−206	.90957	−14	−245	.92783	− 22	−265	.94348	− 26	−265	.95650	− 27	−249	.96705	− 26	−222
165	.91667	0	−160	.93230	−10	−189	.94606	− 16	−203	.95783	− 19	−202	.96761	− 20	−189	.97551	− 19	−167
170	.94444	0	−110	.95492	− 7	−128	.96413	− 11	−137	.97200	− 13	−136	.97851	− 13	−127	.98378	− 12	−112
175	.97222	0	− 56	.97748	− 3	− 65	.98209	− 5	− 69	.98603	− 6	68	.98929	− 7	− 64	.99192	− 6	− 56
180	1.00000	0	−	1.00000	− 0		1.00000	0		1.00000	0		1.00000	0		1.00000	0	

TABLE 4 (Continued)

α	$k=1.2$			$k=1.4$			$k=1.6$			$k=1.8$			$k=2.0$		
	Φ_α	$\delta_\alpha{}^2$	$\delta_k{}^2$	Φ_α	$\delta_\alpha{}^2$	$\delta_k{}^2$	Φ_α	$\delta_\alpha{}^2$	$\delta_k{}^2$	Φ_α	$\delta_\alpha{}^2$	$\delta_k{}^2$	Φ_α	$\delta_\alpha{}^2$	$\delta_k{}^2$
	.00000	0		.00000	0		.00000	0		.00000	0		.00000	0	
5	.06607	− 60	− 20	.07239	− 77	− 24	.07846	− 96	− 27	.08427	−115	− 27	.08981	−137	− 27
10	.13154	−119	− 41	.14401	−151	− 51	.15597	−187	− 55	.15740	−225	− 56	.17827	−266	− 54
15	.19584	−173	− 66	.21413	−219	− 80	.23165	−270	− 86	.24829	−324	− 87	.26409	−382	− 84
20	.25841	−221	− 95	.28207	−279	−113	.30461	−342	−120	.32597	−409	−121	.34613	−479	−117
25	.31879	−263	−129	.34723	−329	−150	.37419	−401	−158	.39959	−476	−158	.42342	−553	−152
30	.37656	−296	−169	.40912	−368	−191	.43979	−444	−200	.46849	−523	−197	.49523	−603	−188
35	.43138	−321	−213	.46735	−396	−236	.50098	−473	−243	.53220	−551	−238	.56104	−630	−226
40	.48302	−337	−261	.52164	−411	−284	.55746	−487	−288	.59042	−561	−280	.62060	−635	−263
45	.53129	−346	−311	.57183	−417	−332	.60910	−487	−332	.64306	−555	−319	.67384	−620	−298
50	.57612	−346	−362	.61788	−412	−379	.65589	−476	−375	.69017	−535	−356	.72092	−590	−328
55	.61751	−341	−412	.65982	−400	−424	.69794	−455	−413	.73195	−505	−387	.76211	−549	−354
60	.65549	−330	−459	.69777	−381	−464	.73545	−428	−446	.76869	−468	−414	.79782	−501	−374
65	.69018	−314	−502	.73191	−358	−499	.76869	−395	−473	.80076	−426	−433	.82852	−449	−387
70	.72174	−296	−539	.76248	−332	−527	.79798	−360	−493	.82859	−381	−446	.85474	−396	−394
75	.75034	−275	−569	.78973	−303	−548	.82368	−324	−506	.85259	−337	−452	.87700	−344	−394
80	.77619	−253	−592	.81395	−274	−562	.84614	−288	−512	.87322	−295	−451	.89580	−296	−389
85	.79951	−231	−607	.83543	−246	−568	.86571	−254	−511	.89090	−255	−445	.91164	−252	−379
90	.82052	−209	−614	.85445	−218	−567	.88275	−221	−503	.90602	−219	−433	.92495	−212	−365
95	.83944	−187	−614	.87129	−192	−560	.89756	−192	−491	.91894	−186	−418	.93614	−178	−347
100	.85649	−167	−607	.88619	−169	−546	.91045	−165	−473	.92999	−158	−398	.94553	−148	−328
105	.87187	−148	−593	.89941	−147	−527	.92169	−141	−452	.93946	−133	−376	.95345	−122	−306
110	.88577	−130	−573	.91115	−127	−504	.93151	−120	−428	.94759	−111	−352	.96013	−100	−284
115	.89837	−114	−548	.92163	−109	−477	.94012	−102	−401	.95461	− 93	−327	.96581	− 82	−261
120	.90983	− 99	−519	.93100	− 94	−447	.94771	− 86	−372	.96070	− 77	−301	.97066	− 67	−238
125	.92029	− 86	−485	.93943	− 80	−414	.95444	− 72	−342	.96601	− 64	−275	.97482	− 55	−216
130	.92990	− 74	−449	.94706	− 68	−380	.96043	− 61	−311	.97068	− 53	−248	.97843	− 45	−194
135	.93876	− 63	−409	.95401	− 57	−344	.96582	− 50	−280	.97482	− 43	−222	.98159	− 36	−172
140	.94699	− 53	−368	.96039	− 48	−307	.97071	− 42	−249	.97853	− 35	−196	.98438	− 29	−151
145	.95469	− 45	−325	.96628	− 40	−270	.97517	− 34	−217	.98188	− 29	−170	.98688	− 24	−130
150	.96194	− 37	−281	.97177	− 32	−232	.97929	− 28	−185	.98494	− 23	−145	.98914	− 19	−111
155	.96881	− 30	−235	.97694	− 26	−193	.98313	− 22	−154	.98777	− 18	−120	.99121	− 15	− 91
160	.97539	− 23	−189	.98185	− 20	−155	.98675	− 17	−123	.99042	− 14	− 95	.99313	− 11	− 72
165	.98174	− 17	−142	.98656	− 15	−116	.99021	− 12	− 92	.99293	− 10	− 71	.99494	− 8	− 54
170	.98792	− 11	95	.99112	− 10	− 77	.99354	− 8	− 61	.99534	− 6	− 47	.99667	− 5	− 36
175	.99399	− 5	− 48	.99558	− 5	− 39	.99679	− 4	− 31	.99769	− 3	− 24	.99835	− 3	− 18
180	1.00000	0		1.00000	0		1.00000	0		1.00000	0		1.00000	0	

TABLE 4 (Continued)

α	$k=2.2$			$k=2.4$			$k=2.6$			$k=2.8$			$k=3.0$		
	Φ_α	δ_{α^2}	δ_{k^2}	Φ_α	δ_{α^2}	δ_{k^2}	Φ_α	δ_{α^2}	δ_{k^2}	Φ_α	δ_{α^2}	δ_{k^2}	Φ_α	δ_{α^2}	δ_{k^2}
	.00000	0		.00000	0		.00000	0		.00000	0		.00000	0	
5	.09509	−159	− 25	.10011	−183	− 23	.10491	−207	− 21	.10949	−233	− 19	.11388	− 260	− 17
10	.18860	−309	− 51	.19842	−354	− 47	.20776	−401	− 43	.21668	−450	− 39	.22519	− 500	− 36
15	.27905	−442	− 79	.29322	−504	− 73	.30666	−569	− 67	.31942	−636	− 61	.33157	− 704	− 55
20	.36513	−551	−110	.38303	−626	−101	.39992	−702	− 93	.41588	−780	− 84	.43100	− 859	− 77
25	.44574	−632	−142	.46664	−713	−131	.48624	−795	−120	.50463	−878	−109	.52193	− 960	− 99
30	.52009	−684	−176	.54319	−766	−162	.56468	−847	−147	.58469	−927	−134	.60336	−1006	−121
35	.58764	−708	−210	.61214	−785	−192	.63472	−859	−174	.65556	−932	−157	.67482	−1002	−141
40	.64816	−705	−242	.67330	−773	−220	.69623	−838	−199	.71718	−899	−178	.73634	− 957	−159
45	.70166	−681	−272	.72676	−738	−245	.74941	−790	−220	.76986	−838	−196	.78834	− 881	−174
50	.74839	−640	−297	.77288	−685	−266	.79472	−724	−236	.81419	−758	−208	.83158	− 786	−183
55	.78873	−588	−317	.81218	−620	−281	.83281	−646	−247	.85097	−667	−216	.86697	− 682	−188
60	.82321	−528	−332	.84528	−549	−290	.86445	−564	−252	.88109	−573	−218	.89554	− 578	−188
65	.85241	−466	−339	.87291	−477	−294	.89045	−482	−252	.90547	−482	−216	.91832	− 479	−184
70	.87696	−404	−341	.89576	−407	−292	.91163	−405	−248	.92501	−399	−209	.93629	− 390	−176
75	.89746	−346	−338	.91453	−342	−286	.92874	−335	−239	94055	−325	−199	.95035	− 312	−165
80	.91449	−292	−329	.92987	−284	−275	.94249	−273	−228	95283	−260	−187	.96128	− 246	−153
85	.92859	−244	−317	.94235	−234	−261	.95349	−221	−214	96248	−207	−173	.96972	− 192	−140
90	.94023	−202	−301	.95249	−190	−246	.96227	−177	−198	97005	−162	−159	.97622	− 148	−127
95	.94985	−166	−284	.96071	−154	−229	.96926	−140	−183	97597	−127	−145	.98123	− 114	−114
100	.95779	−136	−265	.96738	−123	−211	.97484	−111	−167	.98062	− 98	−130	.98509	− 87	−101
105	.96436	−110	−245	.97281	− 99	−193	.97930	− 87	−151	.98428	− 76	−117	.98807	− 66	− 90
110	.96982	− 89	−225	.97724	− 78	−176	.98289	− 68	−136	.98716	− 58	−104	.99039	− 50	− 79
115	.97438	− 72	−205	.98088	− 62	−159	.98578	− 53	−121	.98946	− 45	− 92	.99220	− 38	− 69
120	.97821	− 58	−186	.98390	− 50	−142	.98814	− 42	−108	.99130	− 35	− 81	.99362	− 29	− 61
125	.98146	− 47	−167	.98641	− 39	−127	.99008	− 32	− 95	.99278	− 27	− 71	.99476	− 22	− 53
130	.98423	− 38	−148	.98853	− 31	−112	.99169	− 25	− 84	.99400	− 21	− 62	.99568	− 15	− 46
135	.98662	− 30	−131	.99033	− 25	− 98	.99304	− 20	− 73	.99501	− 16	− 54	.99643	− 13	− 39
140	.98871	− 24	−114	.99189	− 19	− 85	.99419	− 15	− 63	.99586	− 12	− 46	.99706	− 10	− 33
145	.99056	− 19	− 98	.99324	− 15	− 73	.99519	− 12	− 54	.99659	− 9	− 39	.99759	− 7	− 28
150	.99221	− 15	− 83	.99445	− 12	− 61	.99606	− 9	− 45	.99722	− 7	− 33	.99804	− 6	− 23
155	.99372	− 12	− 68	.99553	− 9	− 50	.99684	− 7	− 37	.99777	− 5	− 27	.99844	− 4	− 19
160	.99510	− 9	− 54	.99653	− 7	− 40	.99755	− 5	− 29	.99828	− 4	− 21	.99879	− 3	− 15
165	.99640	− 6	− 40	.99745	− 5	− 29	.99821	− 4	− 21	.99874	− 3	− 15	.99912	− 2	− 11
170	.99763	− 4	− 27	.99833	− 3	− 19	.99882	− 2	− 14	.99918	− 2	− 10	.99942	− 1	− 7
175	.99883	− 2	− 13	.99917	− 2	− 10	.99942	− 1	− 7	.99959	− 1	− 5	.99972	− 1	− 4
180	1.00000	0		1.00000	0		1.00000	0		1.00000	0		1.00000	0	

TABLE 4 (Continued)

α	$k=3.2$			$k=3.4$			$k=3.6$			$k=3.8$			$k=4.0$		
	Φ_α	$\delta_\alpha{}^2$	$\delta_k{}^2$	Φ_α	$\delta_\alpha{}^2$	$\delta_k{}^2$	Φ_α	$\delta_\alpha{}^2$	$\delta_k{}^2$	Φ_α	$\delta_\alpha{}^2$	$\delta_k{}^2$	Φ_α	$\delta_\alpha{}^2$	$\delta_k{}^2$
	.00000	0		.00000	0		.00000	0		.00000	0		.00000	0	
5	.11809	− 287	− 16	.12215	− 316	− 14	.12606	− 345	− 13	.12985	− 375	− 12	.13351	− 406	− 11
10	.23335	− 552	− 32	.24119	− 605	− 29	.24873	− 659	− 27	.25601	− 715	− 24	.26304	− 772	− 22
15	.34317	− 774	− 50	.35427	− 845	− 46	.36490	− 918	− 42	.37513	− 992	− 38	.38497	−1067	− 35
20	.44535	− 940	− 69	.45901	−1021	− 63	.47203	−1103	− 57	.48448	−1185	− 63	.49640	−1268	− 48
25	.53825	−1043	− 89	.55367	−1125	− 81	.56827	−1208	− 74	.58213	−1289	− 68	.59532	−1370	− 62
30	.62083	−1084	−109	.63720	−1160	− 99	.65258	−1235	− 90	.66705	−1307	− 82	.68071	−1378	− 75
35	.69267	−1069	−127	.70924	−1134	−115	.72466	−1195	−104	.73903	−1254	− 94	.75247	−1309	− 86
40	.95390	−1010	−143	.77003	−1059	−128	.78488	−1105	−115	.79857	−1147	−104	.71123	−1184	− 94
45	.80509	− 919	−154	.82029	− 953	−137	.83411	− 982	−122	.84671	−1007	−109	.85821	−1028	− 98
50	.84712	− 810	−161	.86105	− 828	−142	.87356	− 843	−125	.88480	− 853	−111	.89494	− 859	− 98
55	.88108	− 693	−164	.89354	− 699	−143	.90458	− 701	−124	.91437	− 700	−109	.92307	− 696	− 95
60	.90810	− 578	−162	.91904	− 575	−139	.92858	− 568	−120	.93691	− 559	−103	.94421	− 547	− 89
65	.92932	− 472	−156	.93876	− 462	−133	.94687	− 449	−113	.95384	− 435	− 96	.95985	− 419	− 82
70	.94581	− 378	−147	.95384	− 364	−124	.96063	− 348	−104	.96638	− 332	− 87	.97125	− 315	− 73
75	.95849	− 297	−137	.96526	− 281	−113	.97088	− 265	− 93	.97556	− 248	− 77	.97947	− 231	− 64
80	.96818	− 230	−125	.97383	− 215	−102	.97845	− 199	− 83	.98223	− 183	− 68	.98534	− 168	− 55
85	.97554	− 177	−113	.98023	− 162	− 91	.98401	− 147	− 73	.98705	− 133	− 58	.98950	− 120	− 47
90	.98112	− 134	−101	.98500	− 120	− 80	.98807	− 108	− 63	.99051	− 96	− 50	.99244	− 85	− 39
95	.98534	− 101	− 89	.98854	− 89	− 70	.99104	− 78	− 54	.99299	− 68	− 42	.99452	− 60	− 33
100	.98853	− 76	− 78	.99118	− 66	− 61	.99321	− 57	− 47	.99478	− 49	− 36	.99598	− 42	− 27
105	.99096	− 56	− 69	.99315	− 48	− 52	.99481	− 41	− 40	.99607	− 35	− 30	.99702	− 29	− 23
110	.99281	− 42	− 60	.99463	− 35	− 45	.99599	− 29	− 34	.99700	− 24	− 25	.99776	− 20	− 19
115	.99423	− 31	− 52	.99574	− 26	− 39	.99686	− 21	− 29	.99769	− 17	− 21	.99830	− 14	− 16
120	.99534	− 23	− 45	.99660	− 19	− 33	.99752	− 15	− 24	.99819	− 12	− 18	.99869	− 10	− 13
125	.99621	− 17	− 39	.99726	− 14	− 28	.99802	− 11	− 21	.99857	− 9	− 51	.99897	− 7	− 11
130	.99690	− 13	− 33	.99778	− 10	− 24	.99841	− 8	− 18	.99886	− 6	− 13	.99919	− 5	− 9
135	.99746	− 10	− 28	.99819	− 8	− 21	.99872	− 6	− 15	.99909	− 5	− 11	.99936	− 4	− 8
140	.99791	− 7	− 24	.99853	− 6	− 17	.99896	− 4	− 12	.99927	− 3	− 9	.99948	− 3	− 6
145	.99830	− 6	− 20	.99880	− 4	− 14	.99916	− 3	− 10	.99941	− 2	− 7	.99959	− 2	− 5
150	.99862	− 4	− 17	.99904	− 3	− 12	.99933	− 2	− 8	.99953	− 2	− 6	.99967	− 1	− 4
155	.99891	− 3	− 14	.99924	− 2	− 10	.99947	− 2	− 7	.99963	− 1	− 5	.99974	− 1	− 3
160	.99916	− 2	− 11	.99941	− 2	− 8	.99959	− 1	− 5	.99972	− 1	− 4	.99980	− 1	− 3
165	.99939	− 2	− 8	.99957	− 1	− 6	.99970	− 1	− 4	.99980	− 1	− 3	.99986	0	− 2
170	.99960	− 1	− 5	.99972	− 1	− 4	.99981	− 1	− 3	.99987	0	− 2	.99991	0	− 1
175	.99980	0	− 3	.99986	0	− 2	.99991	0	− 1	.99993	0	− 1	.99995	0	− 1
180	1.00000	0		1.00000	0		1.00000	0		1.00000	0		1.00000	0	

control gross errors in the fourth. The modification, rounding, checking, and printing then proceeded as with the 7 decimal table, the 5 being added, of course, in the sixth place.

4. SUMMARY

The mean direction α_0 and the vector strength $\bar{a}$ of circular data as used by climatologists are calculated from (1.2′) and (1.3). The same statistics may be used for data other than weather. With data grouped in monthly class intervals the calculations of these statistics are facilitated by (1.4) and (1.5). The assumption that the maximum likelihood estimate of a "true value" shall be given by the mean direction leads to the circular normal distribution (2.1). The central location parameter of this distribution is estimated by the mean direction α_0. The parameter k is estimated from $\bar{a}$ by means of Table 2. The radii for plotting the distribution in aequiareal polar scale are obtained by multiplying the values given in Table 3 by $\sqrt{n/12}$, where n stands for the total frequency. Areas of the distribution are given in Table 4. Numerical examples will be given in a forthcoming article.

ACKNOWLEDGMENTS

The authors are indebted to Arnold Court for drawing their attention to the problem and for suggesting the aequiareal procedure; to Dr. Churchill Eisenhart and Dr. J. C. P. Miller, of the National Bureau of Standards, and to Dr. Hans G. Hertz, of the Naval Observatory, for assistance in locating convenient tables of the Bessel functions $I_n(x)$; to the Manhattan Life Insurance Company and the National Bureau of Economic Research for the use of calculating equipment; and to Mario Bosso, W. Braddock Hickman, Donald Ladd and Andrew Scaglia for assistance with machine difficulties.

REFERENCES

[1] Arnold, K. J., "Spherical Probability Distributions," Dissertation, Massachusetts Institute of Technology, 1941.

[2] Comrie, L. J., "The application of the Hollerith tabulating machine to Brown's tables of the moon," Royal Astronomical Society, *Monthly Notices*, 92 (1932), 694–703.

[3] Comrie, L. J., "The application of commercial calculating machines to scientific computing," *Mathematical Tables and Other Aids to Computation*, 2 (1946), 149–59.

[4] Frullani, G., "Sopra la dipendenza fra i differenziali delle funzioni e gli Integrali definiti," *Memorie della Società italiana* (Modena), 18 (1820), 458–517.

[5] Gray, A., Mathews, G. B., and Mac Robert, T. M., *A Treatise on Bessel Functions*, 1922.

[6] Gumbel, E. J., "Die Bahnen der Betastrahlen," *Zeitschrift für angewandte Mathematik und Mechanik* (Berlin), 3 (1923), 320.
[7] Legendre, A. M., *Exercices de Calcul Intégral*, Paris, 1816.
[8] Leighly, J. B., "Graphic studies in climatology: the polar form of diagram in the plotting of the annual climatic cycle," *University of California Publications in Geography* (Berkeley), 2 (1928), 387–407.
[9] Lévy, P., "Sur l'arithmétique des lois de probabilités enroulées," *Société Mathématique de France, Comptes Rendus*, 1938, p. 32.
[10] Marcinkiewicz, J. M., *ibidem*, p. 34. (See also remarks by Lévy, p. 36)
[11] Mises, R. v., "Über die 'Ganzzahligkeit' der Atomgewichte und verwandte Fragen," *Physikalische Zeitschrift*, 19 (1918), 490–500.
[12] Pearson, Karl, ed., *Tables for Statisticians and Biometricians*, Part II, 1931.
[13] Pearson, Karl, ed., *Tables of the Incomplete Beta-Function*, Cambridge, 1934.
[14] Pólya, G., "Zwei Aufgaben aus der Wahrscheinlichkeitsrechnung," *Vierteljahrsschrift der Naturforschende Gesellschaft* (Zurich), 80 (1935), 123–30.
[15] Pólya, G., "Sur quelques points de la théorie des Probabilités," Paris, Université, Faculté des Sciences, Institute Henri Poincaré, *Annales*, Vol. 1 (26 August, 1930), 117–61.
[16] Pólya, G., "Zur Statistik der sphaerischen Verteilung der Fixsterne," *Astronomische Nachrichten*, 208 (1919), 175–80.
[16a] Royal Society, London, *British Association Mathematical Tables*, 6, 1937.
[17] Royal Society, London, *British Association Mathematical Tables*, 10, in press.
[18] Schelling, H. v., "Most frequent particle paths in a plane," *Transactions of the American Geophysical Union*, 32 (1951), 222–6.
[19] U. S. Department of Commerce, National Bureau of Standards, Applied Mathematics Series 5: *Tables of the Sines and Cosines to Fifteen Decimal Places at Hundredths of a Degree*, 1949.
[20] Wintner, A., "On the shape of the angular case of Cauchy's distribution curves," *Annals of Mathematical Statistics*, 18 (1947), 589–93.

BOOK REVIEWS

Mathematics Essential for Elementary Statistics. Revised Edition. *Helen M. Walker* (Professor of Education, Teachers College, Columbia University). New York: Henry Holt and Company, 1951. Pp. xiii, 382. $2.75.

LEO A. GOODMAN, *University of Chicago*

THIS book, a self-teaching manual for the adult who has forgotten or never grasped elementary mathematics, still remains without serious competitors.

The first edition (1934) of this book was reviewed by Herbert Toops in this *Journal*, Volume 30 (1935), pp. 483–84. This revised edition contains new sections dealing with simple operations with common and decimal fractions, inequalities, linear interpolation, permutations and combinations, the binomial expansion, the multinomial expansion, ratios and proportions, trigonometric identities, second degree equations, degrees of freedom, and weights and measures. Many of the explanations appearing in the 1934 edition have been expanded. In order to keep the book small (the first edition contained 246 pages) all those sections of the first version which related specifically to the content of statistical method have been eliminated.

Every topic discussed is accompanied by self-scoring exercises, answers for which are provided in a key at the back of the book. Some of these answers (e.g., p. 9 problem 14, p. 34 problem 11 (2), p. 46 problems 4 and 5, and others) are incorrect, but this should not confuse the student who has mastered the topic.

Elements of Statistics. *C. G. Lambe.* London: Longmans, Green and Company, 1952. Pp. viii, 112.

CARL F. KOSSACK, *Purdue University*

THE author sets forth the object of this text as follows: "This work is written primarily for the use of engineering students, but it is hoped that it will be a useful introduction to the Theory of Statistics for scientists and experimentalists of all kinds." Such a goal is a very ambitious one and one which has interested the reviewer for some time. It embraces the difficult question of whether it is necessary to have separate introductory courses in statistics for each major area or if a single first course will suffice. Since this text also includes a section on Quality Control, one finds himself faced with the further problem of whether this phase of statistics can be combined with other phases or needs to be treated in an independent course. Unfortunately, the text being reviewed does not help much if one is interested in answering such questions.

In fact, the text is so brief, consisting of only 104 small pages, that one feels that many of its weaknesses can be attributed to this limitation in space. The material covered falls into three major catagories: elementary

statistical methods, the theory of errors, and quality control. The approach lacks inspiration. It follows the usual outline: frequency distributions, measures of dispersion, continuous distributions, etc., each chapter being very restricted in its nature, consisting of around ten pages including the exercises. The large number of exercises associated with each chapter is to be commended, but one cannot help but feel that they have been introduced either as arithmetic projects or for mathematical gymnastics rather than to develop the student's appreciation of statistics.

The chapter on the theory of errors is the most worthwhile one in the book. The material in this chapter deserves being considered for inclusion in more courses in statistics than it is today. It is apparent that the eight pages used by the author to discuss quality control are not enough to do the subject justice.

One could mention particular difficulties encountered in reading certain sections. Even the first sentence in the book gives rise to some question, "Statistics may be defined as the scientific treatment of numerical data." It is apparent that the statistician will soon inherit the earth. Not enough attention is given to explaining the real purpose behind the many statistical ideas and measures that are introduced. For example, the concept of independence of two variables is never really explained, let alone defined, although it is used.

An improvement in the text could be achieved if more space and attention were given to the sections on the theory of errors and quality control, with the other material restricted to an introductory section stressing more the basic concepts of statistics, especially those needed to appreciate the work covered in these two areas.

Introducción a los métodos de la estadística. 1ª Parte. *Sixto Rios.* Madrid: 1952. Pp. 205. Paper.

PAUL R. HALMOS, *University of Chicago*

THIS is a charming and elementary book that fulfills, within the limits the author sets for it, the promise of the title. The mathematical level of the book is not advanced; nothing more high-powered than integral calculus is ever used. The topics discussed are: statistical tables, graphical representation, frequency and probability, sampling, random variables, multidimensional statistics, correlation, regression, sampling distributions, and tests of hypotheses. Although at times the definitions are a little vague (cf. the definitions of statistical variable and random variable on pp. 35 and 36), the large number of examples, complete with detailed tables and graphs, is likely to give the reader a sound intuitive grasp of the subject. A notable feature of the book (emphasized by Herman Wold in his preface) is the number, variety, and interest of the exercises. The reader's statistical imagination will probably be stimulated when he is asked whether meat consumption in Madrid can be adequately studied by using samples of telephone subscribers

(p. 8), and when he is requested to perform various intricate coin-tossing experiments that illustrate the possible presence of correlation in multi-dimensional situations (p. 97). As a pre-technical introduction, with the purpose of telling the experimenter what ideas he is likely to encounter in his study of statistics and what statistics is likely to do for him, the book is highly recommended.

Statistik Tidskrift. Nr 1, 1952. Stockholm: Statistiska Centralbyrån, Januari 1952. Pp. 83. Paper.

STATISTIK Tidskrift published by the Central Bureau of Statistics of Sweden is a new journal whose main purpose is to give summaries of current statistics at an earlier date than official statistical reports usually do. It will also contain articles on new statistical techniques, international statistical collaboration, reviews of books in statistics, and so on.

Among the articles in this first issue the following two are of special interest to statisticians, in general: one by H. Cramér on statistical methodology and the other by T. Dalenius on operations analysis.

U. G.

A Guide to Tables of the Normal Probability Integral, National Bureau of Standards Applied Mathematics Series 21, Washington, U.S. Government Printing Office, 1952, 15 cents. Pp. iv, 16.

THIS Guide" its Introduction says, "makes available in one place certain information that will be helpful in utilizing the normal probability tables given in standard statistical texts and other important sources. . . . The aim has not been to present an exhaustive or historical compilation but rather to include the more important sources likely to be accessible in libraries of American universities, research organizations, etc., as well as some of the more recent publications that educational institutions, research organizations, and students of statistics might be encouraged to acquire." There are 62 entries in the bibliography.

The tables are classified into two main categories, *direct* and *inverse*, according to whether the probability is shown as a function of the unit normal deviate, or the deviate as a function of the probability. Each category is divided into five groups, according to the kind of probability tabulated: central, semi-central, two-tail, single-tail, or cumulative. "Parts I and II tell which sources give tables for the indicated types of segments of area under the normal curve and give the essential characteristics of the tables. These features are succeeded by a description of interpolation methods (Part III) especially applicable to certain of the tables mentioned. Then follow two brief comments [i.e., two long sentences] . . . giving the relationship between some types of normal tables" and the incomplete gamma and chi-square tables.

W. A. W.

Consumers' Price Index. Report of a Special Subcommittee of the Committee on Education and Labor Pursuant to H. Res. 73, *Subcommittee Report No. 2*, House of Representatives, 82nd Congress, 1st Session, 1951, 39 pp.

JULES BACKMAN, *New York University*

EARLY in 1951, the United Electrical, Radio, and Machine Workers of America published a study containing extensive criticisms of the Consumers' Price Index. This study followed the pattern of the Meany-Thomas Report which was used to denounce the Bureau of Labor Statistics and the Consumers' Price Index during World War II. To evaluate these charges and to determine the present adequacy of the Consumers' Price Index, a special Subcommittee of the House Committee on Education and Labor was established. After hearing a number of witnesses, this Committee prepared the report under review.

This report is an excellent brief survey of how the Consumers' Price Index is constructed, its use, and its history. The Committee found that the UE charges were without merit and that the Consumers' Price Index "is an excellent index." The Committee made several recommendations which appear to be solidly based. Thus, for example, they recommended that the BLS make periodic studies of "changes in living costs due to changes in non-price factors" and "the extent of changes in prices paid by wage earners and lower salaried clerical workers who moved from one community to another." Such studies would be useful particularly during periods characterized by major upheavals in the economy such as those which attend an armament program. The Committee properly suggested that the Consumers' Price Index should not incorporate these types of changes.

The Committee also reviewed the suggestion that income taxes should be included in the index. It recommended a continuation of the present practice under which the BLS includes excise and sales taxes, but excludes income taxes. All citizens have a responsibility to pay their fair share of the cost of government. If income taxes were included in the Consumers' Price Index, the net result would be to exempt from such responsibility all workers who are covered by escalator clauses or who received wage increases based upon changes in the Consumers' Price Index. The adoption of a general sales tax, however, would create serious problems in this connection and would require a reexamination of the BLS method of handling such taxes.

The Committee also recommended that BLS should follow "a policy of keeping the index continuously under review, with revisions as required when important changes in buying habits occur, rather than infrequent complete revisions as has been the practice in the past." If such a continuous review were made, there would be little need for the major revisions of the index such as the Bureau is now engaged in making. The Committee properly suggests that BLS "should be extremely cautious against making too frequent changes of the index which are minor in character."

The Committee has recognized that to effectuate the changes it suggests would require adequate financial support for the Bureau. This in itself is a contribution since very often such committees make recommendations

without being willing to face the fact that their execution would cost money. It is very healthy to have periodic reviews of an index as important as the CPI. The Steed Committee is to be congratulated for a very workmanlike job.

A Dollar Index of Soviet Machinery Output, 1927–28 to 1937. *Alexander Gerschenkron*, assisted by *Alexander Erlich*. Santa Monica: The Rand Corporation, April 6, 1951. Pp. v, 352. Paper. Not published for sale; a limited number of copies are available on request from interested scholars.

JOHN CRAWFORD, *United Nations*

THIS study is mainly devoted to the construction and appraisal of a new index based on the physical quantities of 128 items of machinery produced in Russia between 1927–28 and 1937 weighted with 1939 U. S. factory prices. This unusual index is described by the author as "part of a comprehensive recomputation of the rate of growth of industrial output in Russia . . . justified by the manifold deficiencies of the official Soviet index of industrial production."

Briefly, it is charged that the official Soviet indexes have grossly overstated the real rate of growth of industrial output. The machinery index is regarded as exemplifying this tendency since it is felt to reflect in exaggerated form the main weakness of the Soviet indexes, namely, the introduction of new commodities at approximately their current prices during the period of general inflation subsequent to 1926–27, the weight year of the indexes. This results in an upward bias because the output of the new items expanded faster than the old. Thus an examination of the index computed by Gerschenkron reveals that 24 items were first included in 1931 and 21 in 1932. The value of these items in 1937 (at 1939 U. S. prices) was 3.5 to 4 times their value in the year they were first included, whereas the total index only doubled during the same period.

As evidence of the charge that the "1926–27 prices" of machinery are inflated (presumably in the sense that they contain prices of later years for commodities not produced in 1926–27), the author refers to data presented in the Soviet Plan for 1941. This source contains values of planned output for 1941 at current and at 1926–27 prices for a number of different Ministries. These indicate that the 1941 planned output at current prices for the electrical and other machinery industries exceeds that in 1926–27 prices by only 10 per cent, whereas 1941 values for the ferrous metal, coal, and woodworking industries are more than double the values in 1926–27 prices. It is concluded that the similarity of values at 1941 and 1926–27 prices for the machinery industries (in which the increase in physical production is known to have been large) is explained by the fact that so-called 1926–27 prices contain prices of subsequent years for commodities not produced in 1926–27. The author rejects as a sufficient explanation cost reductions in the machinery industry (said to have amounted to 45 percent over the period of the Second Five-year Plan, 1933–1937) on the ground that this reduction was insufficient as well as on the basis of statements of Soviet writers.

In addition to the effect of the general price inflation, the weights are believed to be inflated also because the prices used for the new commodities are said to be those which prevailed in the initial period of mass production before the usual increases in productivity took place.

Aside from the problem of new commodities, the use of 1926–27 prices as weights is considered to have imparted an upward bias to the index because increases in output and productivity are positively correlated. Thus the use of beginning-period prices as weights would result in an index showing a larger rate of growth than the use of end-of-period prices. This proposition is tested with U. S. price and quantity data for machinery items common to the years 1899 and 1909, 1899 and 1923, 1909 and 1923, and 1909 and 1939: indexes computed with beginning-period prices as weights show from two to seven times as large an increase as when end-of-period prices are used. The longer the period of comparison the greater is the relative increase.

To get around the problem caused by the "process of changing scarcity relationships" involved in the economic development of a country, and lacking a consistent set of Soviet prices for machinery, the author used 1939 U. S. price relationships. His choice of U. S. prices is defended on the grounds of their availability and the fact that their use permits a comparison of the absolute levels of U. S. and Soviet machinery production. The prices used were drawn from Census of Manufactures data and from American businessmen, who having dealt with Russia were able to identify and price American counterparts of Soviet machinery items. Use of the Census limited the period chosen to census years and from these 1939 was selected as being the most recent available for the purpose, data for 1947 being regarded as distorted by the war. According to the author, "The problem of using as index weights prices of a different country is essentially the same as that of basing an index of a country's output over a given period of industrialization alternatively on prices pertaining to the beginning and to the end of a period." It would seem, then, that if the Russian machinery industry during the 1930's is regarded as having been at the same stage of economic development as, for instance, that in the U. S. just prior to the first World War, the use of 1914 U. S. prices rather than 1939 would have more nearly approximated Soviet scarcity relationships.

The fact that in 1937 the Soviet index exceeded 3.2 times the index computed by Gerschenkron (1927–28 = 100) is explained by (1) an unknown amount of downward bias in the latter measure; (2) pricing errors of the Soviet index; and most important of all, (3) the effects of using as weights for the computed index prices of a year (1939) that approximate the end of the period (1937) and are prices of a more highly developed economy. As the author states, "The use of American prices of the year 1939 is an altogether *sui generis* way of looking at Soviet Output. The index of Computed Output, therefore, is *not* a gauge by which the Soviet Indexes can be adjusted."

Accordingly, the charge of bias in the Russian index still seems to rest on the validity of these propositions: that there was a substantial degree of inflation in Soviet machinery prices after 1927–28, and that in a period of

economic development the use of base period weights leads to a higher index than weights based on end-of-period prices.

Whether the reader will agree with the author's conclusion that "with all its deficiencies the [Gerschenkron] index should still provide a possible basis for gauging the development of a crucial branch of Russia's industrial economy" will depend on the meaning attached to measures of industrial production. Aside from the pricing deficiencies of the Soviet index the question of the choice of beginning- or end-of-period weights must rest on the use of the indexes. As to the use of U. S. prices as weights, it would appear that the meaning of a production measure is anchored in its economic setting so that whatever phenomenon the new index can be said to measure, it is not the physical volume of Russian machinery output.

Students of production indexes will be indebted to Professor Gerschenkron not only for the general clarity of his description but also for the 288 pages of appendices which set forth the basic data and computations underlying the study. May the Russians take heed.

Business Fluctuations. *Robert A. Gordon.* New York: Harper and Brothers, 1952. Pp. xvi, 624. $5.00.

HENRY H. VILLARD, *The City College, New York*

PROFESSOR Gordon has written an able and unusual volume on business fluctuations, placing main emphasis on actual cycles and minimizing the space devoted to cycle theory. The material is divided into three parts: the first, covering background understanding and tools of analysis; the second, the nature and causes of fluctuations; and the third, prediction and control.

Part I starts with a discussion of the income and transaction versions of the quantity equation, and of national income, gross national product, and related concepts. Next the determinants of aggregate demand are considered. The relationship between consumption and investment is examined at length, both in theoretical, *ex ante* terms and in the light of statistical, *ex post* evidence, and the theory of the multiplier is analyzed. The first part ends with a discussion of fluctuations in investment and the role of the acceleration principle. The conclusion is that the acceleration principle—either as an explanation of investment fluctuations or in conjunction with the multiplier as an explanation of entire cycles—is a weak reed without noticeable empirical support, and that the Keynesian determinants via liquidity preference and the supply of money are of no more practical use. The detailed subsequent discussion is therefore necessary if fluctuations are to be understood.

The second part, on the nature and causes of fluctuations, constitutes the core of the volume. It starts with a presentation of the meaning and available measures of business activity, including aggregative measures such as the GNP, bank debits, and employment and index numbers such as those of the Federal Reserve and the American Telephone Company. Next the char-

acteristics of private-enterprise economies making for secular and cyclical change are set forth, with population growth, technological developments, alterations in tastes, "structural changes," and wars placed in the former category and production for profit, price-cost relationships, and the use of unstable money in the latter. Stress is laid on the inevitable tendency toward fluctuations under these conditions, which the much emphasized relationship between saving and investment summarizes but does not explain. The kinds of economic change which result are set forth, including seasonal, secular, cyclical, and irregular movements. The nature and past behavior of secular movements is covered, together with Kondratieff long waves (which are considered to be unconfirmed) and building cycles (which are considered significant in connection with business cycles of unusual severity). This clears the way for a presentation of how the economy behaves during business cycles, which relies heavily upon the studies of the National Bureau of Economic Research (whose method of measuring the cycle is described at this point), and leads in turn to a discussion of how the economy generates business cycles. The most original part of the presentation is a welcome distinction between minor and major depressions: the former are likely to result from over-expansion of inventories or overproduction in particular areas and be ended by liquidation of inventories or absorption of the specific excess capacity; the latter are likely to result from impairment of long-term investment opportunities and be ended by accumulated needs to replace worn-out capital and renewed expansion of growth industries, with building activity sometimes playing an important initiating role.

Two of the last four chapters of Part II are devoted to business cycle theories and two to a description of American economic developments from 1919 to the Korean War. The theorists discussed are grouped as "business-economy" theorists (Mitchell, Veblen, Pigou, and Metzler), monetary theorists (Hawtrey, Wicksell, and Warburton), a "capital-shortage" group (Hayek, Spiethoff, and Cassel), a "partial-overinvestment" group emphasizing the exhaustion of specific investment opportunities (Schumpeter, Hansen, and Robertson), and a group stressing overinvestment in relation to final demand (Clark, Harrod, Kalecki, and Hicks). Underconsumptionists and weather theorists are mentioned briefly, while an appendix is devoted to econometric model-builders. The final two chapters describing actual fluctuations in detail are in many ways the most original and interesting in the book. That they are urgently needed is obvious, now that so many students have no recollection of anything except the prosperity that has prevailed since World War II. The presentation contains many suggestions as to the factors responsible for the fluctuations that occurred without sacrificing balance in the description of the events themselves.

The final section of the volume is devoted to prediction and control. Two broad methods of forecasting business activity are distinguished: historical patterns and "cross-section analysis." In the first group only the statistical indicators of the National Bureau are found to be promising, but various methods of projecting the GNP and its components (and relating them to

an industry and ultimately to a firm) are found to be worthwhile. International aspects of instability are also discussed in this section, including the way in which fluctuations are transmitted and the measures needed to reconcile internal and external stability. But the bulk of the section is devoted to the achievement of domestic stability, including not only full employment but also the elimination of significant price changes. The extent to which these objectives can be realized by controlling consumption and private investment, by using monetary and fiscal policies, and by influencing wages and prices are discussed at length. A final chapter summarizes the extent to which governments in the United States and elsewhere are today formally committed to achieve full employment and the actual preparations that have been made. In attaining stability major importance is attached to fiscal policy, but the broad setting in which such policy must operate and the way in which other policies can be used to help is also made clear. Throughout, the emphasis on how to achieve stability is unusual and will be welcome to those whose students insist on knowing what can be done about economic fluctuations.

Professor Gordon's volume is a solid and substantial achievement. If I have any reservation, it is that the book is a bit too solid for its original objective of providing a text for an upper-division course required of all business administration majors at the University of California. Almost all texts dealing with business cycle theories have felt an obligation to include theories which have been discredited professionally; Professor Gordon seems to have felt a similar obligation to include almost everything that business cycle economists have been doing in recent years, whether or not the results have been particularly fruitful. The result is a book with quite unequalled references to the periodical literature which should be of unique value to the graduate students just starting in or to professional economists who have not kept up with recent developments. But I believe—though it is clearly a matter of judgment —that the volume would have been of greater value to undergraduates if perhaps a third or so of the material had been omitted and more emphasis placed on crucial areas of analysis and policy.

To illustrate: most of the tools considered in the first section are never subsequently used to justify the space (and student energy) devoted to their development. Again in Part II I believe that the 140 pages leading up to how cycles are generated might well have been condensed in the interest of an elaboration of the forces involved, which are presented rather sketchily. I feel, for example, that too little stress is laid on the increasing *postponability* (with little sacrifice of current satisfaction or production) of expenditures on consumer and producer durables, investment, and replacement at the end of a period of prosperity, and on the increasing *inability to postpone* expenditures (without immediate repercussions on satisfaction or production) after an extended depression. The discussion of fiscal versus monetary policy in Part III also seems to me to reflect an unawareness—admittedly one widely prevalent at the present time—of the fact that fiscal policy is generally subject to the same limitations as monetary policy in checking

an inflation when a commitment to stabilize government bonds at any particular level has been accepted—except perhaps to the extent that a budget surplus increases saving and so permits a lower level of interest rates to be maintained without inflation. Certainly, holding of the proceeds of a surplus in idle Treasury deposits, which Gordon recognizes is usually necessary if the surplus is to be fully deflationary during a period of prosperity, will in most cases have the same repercussions on the government bond market as a similar reduction of deposits in the hands of the public through the use of monetary policy. Finally, I wish that Gordon had discussed price stability somewhat more realistically. Despite a preference for *constant* prices in the long run (p. 487) and a suggestion that stabilization may require "some sort of supervision" over wage and price policies (p. 498), the subsequent detailed discussion (pp. 551–558) offers little except education of labor leaders and the possibility of letting unemployment rise as high as 6 or 7 per cent. Actually constant prices mean that average wages, including fringe benefits, cannot increase as much as a nickel an hour a year.[1] Given such facts as the steady increase in wages from 1933 to 1937, when unemployment was far higher than 7 per cent, and Mr. Lewis's recent characterization of an effort to cut his miners back to a wage increase of rather more than 10 per cent as exhibiting a sadistic "penchant for robbing miners' babies of life-giving milk," I feel that more discussion of just how labor leaders are to be educated would be of educational value to all of us!

But these are largely matters of preference and emphasis. The sort of book I would like for undergraduates doesn't exist, and, the proof of the pudding being in the eating, I shall be using Gordon's volume in my undergraduate business cycles course next semester.

Private and Public Investment in Canada, 1926–1951. *O. J. Firestone.* Ottawa: Department of Trade and Commerce, 1951. 4°. Pp. 254. Paper.

MILLARD HASTAY, *National Bureau of Economic Research*

THERE should be something of value in this complex and formidable monograph for every serious student of investment.

In origin, it is a source book of Canadian fixed capital expenditures in the period 1926–51; and this aspect of the work finds expression in Part II, which consists of 64 quarto-size pages of tabular material on Canadian investment flows, and Part III, which comprises 38 pages of explanatory material on concepts and definitions, sources and estimating techniques, and the quality of the estimates. The impressive thing about this side of the work is its scale. The estimates cover all outlays on durable productive assets by every sector of the economy. This means not only gross investment, but also repairs and maintenance, and not only by the business community, but also by such institutions as churches, schools, universities and hospitals, by governments both on durable physical assets and on development and con-

[1] Assuming a 2 per cent annual increase in productivity per man hour and no redistribution of income or increase in unemployment. As dividends are about 5 per cent of wages and salaries, opportunities for redistribution are severely limited.

servation of natural resources, and by individuals on housing. The figures are shown separately for the private and public sectors, and also for industrial groupings based on the Canadian Standard Industrial Classification, defined to include publicly as well as privately owned establishments. Finally, both gross investment and repairs and maintenance are divided into outlays for construction and outlays for machinery and equipment. Most of these distinctions can, in turn, be crossed; and several of them, such as "the public sector" or industrial groupings such as "manufacturing" and "utilities," are themselves multiple classifications. The end product is a set of 123 tables—each containing, on the average, four non-duplicating time series—which presents a picture of Canadian investment that cannot be matched either in comprehensiveness or in detail by published estimates covering the United States. These estimates are of course of varying quality, and every user must assure himself that their defects are not critical for his purposes. But they have been put together on a systematic plan and appear from the descriptions to be at least as reliable as the best comparable estimates compiled for the United States.

More than half the book is taken up by Part I, which constitutes a 133-page interpretative summary of the stark facts contained in the statistical sections. This summary is comprehensive rather than intensive, but it adds up to a highly instructive account of the growth and changing structure of the Canadian economy as disclosed in a quarter-century of investment experience. Every reader of this summary will be impressed by the author's wide knowledge of Canadian economic development and by his sure sense of the special factors which serve to diversify investment patterns in different regions and industries. And if the reader is familiar with United States experience, his attention will be arrested by a number of striking parallels. For the diversity of Canadian resources is scarcely less than that of our own, and the Canadian people are embarked on a course of industrialization which recapitulates much of our economic history at a substantially stepped up pace. Thus, in the last quarter century, manufacturing has gone ahead of agriculture in the number of persons employed and now contributes roughly three times as much to national income. Utilities, too, have undergone a phenomenal expansion until, in relation to population, Canada's railroad, telephone, and electric power networks rank among the largest and most advanced in the world. And, finally, Canada has pioneered in working out viable relationships between free private enterprise and a federal system of government. In all of these respects Canadian experience supplements our own, and Dr. Firestone's monograph should find a grateful audience among students of investment in this country.

Psychological Analysis of Economic Behavior. *George Katona.* New York: McGraw-Hill Book Company, 1951. Pp. ix, 347. $5.00.

MARGARET G. REID, *University of Chicago*

"THE author has set for himself the task of describing a psychological approach to economic analysis and the current research in the field of

economic behavior," to be done in such a manner as to reach "the general public as well as the expert." (p. v.) To the reviewer the book has more of the earmarks of a text for undergraduate students than a monograph for experts, or of a book for lay readers, however interested they may be "in present-day American economic life." Furthermore it appears to deal with specialized research rather than what can be described as "the current research."

The book is divided into five parts: Part I. Problems and Tools; Part II. Consumer Behavior; Part III. Business Behavior; Part IV. Economic Fluctuations; and Part V. Research Methods. Part II contains more than one third of the pages, and almost all the statistical data presented about economic behavior. Parts III and IV deal with concepts, theories and generalizations based on non-statistical observations, despite the author's judgment that "statements not susceptible of empirical validation have no place in psychology." (p. 29.) Unsubstantiated and in fact highly dubious statements occur; for example, "Business and labor leaders have become increasingly aware of their responsibility to the public as well as of the fact that their own interests can best be promoted by striving for what is in the general interest." (p. 294.) One chapter in Part I is devoted to "What kind of psychology?" In this the student is presumably to be introduced to the "solidly established tenets of modern psychology." (p. 8.) However, the psychological ideas actually used in later discussion are so much a matter of common knowledge that the person who omits this chapter is not likely to feel handicapped.

Part V dealing with Research Methods comprises 34 pages. Attention centers on the investigation of attitudes rather than of behavior as such and types of schedules and questions and their use in getting reports on attitudes. The text ends with two pages devoted to "research projects" in which the author emphasizes the need for examining in so far as possible the effect of one factor at a time. He expresses the belief that considerable progress has been made in this direction. He is, however, very vague or silent as to the extent to which a systematic field of knowledge has developed in the psychology of understanding economic behavior or the concrete gaps that exist and should be a challenge to research workers.

With minor exceptions data presented in Part II are those provided by the Consumer Finance Surveys conducted by the Survey Research Center of the University of Michigan for the Federal Reserve Board. Out of 34 formal tables 32 present data from this source, one describes the components of the national budget in selected years and one presents hypothetical data. These tables furnish a useful digest of the findings of the Consumer Finance Surveys and include some data not previously published.

The author states at the outset that "it will be shown in the book that studying the motives, attitudes, and expectations of consumers and businessmen contributes to the understanding of spending, saving, and investing. . . . The results of psychological-economic studies will supplement the traditional analysis of supply, demand, income and consumption." (p. 4.)

The expectations raised by considerable emphasis at the outset on "motives, attitudes and expectations" are likely to be disappointed for many readers. Out of 32 tables presenting findings from the Consumer Finance Surveys only 11 include data on attitudes, motives and expectations. The author does not, however, devote much space to considering the unique contribution of the two types of data. He does emphasize that people are able to predict change in future income but little or no recognition is given to the fact that the incomes of individuals and of families tend to move through well-defined cycles so that individuals tend to be able to predict the course of their income even when unable to predict the course of national income. The author does state "Perhaps primarily, the analyst can rely on the results of cross tabulations between forms of behavior and certain characteristics of people." (p. 75.) The context would seem to imply that the author had in mind things other than "motives, attitudes and expectations," although contact with certain sources of information is included in characteristics. The data used have a uniqueness with respect to characteristics of people that for many will be the major attraction of this book. Whereas the earlier analyses of consumption and current savings in relation to economic levels were largely confined to the effect of income of one year the analysis presented here in addition includes income of two years, extent of change in income and assets and especially in liquid assets. The author is, however, handicapped in his analysis of numerous variables by the necessity of using a small national sample.

Because of the excessive attention to variations in behavior factors among individuals readers may feel that they lose sight of a major fact of interest—namely the mass response to certain types of change. Some of the problems of aggregation are touched on but only in a rather tentative fashion.

The book is marred by occasional carelessly worded or superfluous statements. Some confusion, for example, exists in the meaning of "fact" and "knowledge" and "response" of interviewees. It would hardly seem necessary to make an extensive investigation to disclose that "people's motives, or needs, desires and hopes . . . were conflicting." (p. 73). Nor does it seem necessary to hypothecate after the accumulated observation of human history, not to mention the development of philosophy, that "some of our expenditure may be found to be nothing but expressions of whims or the results of emotions." (p. 69.) In view of the classification of money spent on durable goods as expenditures the reader is likely to be a little puzzled at the statement that "the net sum of those relatively infrequent outlays which are not used for the satisfaction of immediate needs is considered saving." (p. 151.) We are also told that "There were people who felt that they were seriously affected by the 'depression'—mostly people with low incomes." (p. 281.) One is left to wonder whether they were people with low incomes because of the depression. After a general discussion of "probable reactions to policy" without any empirical evidence the author concludes that "it is often possible to differentiate between appropriate and inappropriate responses to economic policy and to economic change." (p. 288.)

PUBLICATIONS RECEIVED

Anderson, R. L., and Bancroft, T. A. *Statistical Theory in Research.* New York: McGraw-Hill Book Co., 1952. $7.00.

Andrews, F. Emerson. *Corporation Giving.* New York: Russell Sage Foundation, 1952. $4.50.

Ashby, W. Ross. *Design for a Brain.* New York: John Wiley and Sons, 1952. $6.00.

Batchelor, James H. *Operations Research, A Preliminary Annotated Bibliography.* Cleveland: Case Institute of Technology, 1952. Paper.

Belcher, John C., and Sharp, Emmit F. *A Short Scale for Measuring Farm Family Level of Living: A Modification of Sewell's Socio-Economic Scale.* Stillwater: Oklahoma Agricultural Experiment Station, 1952. Paper.

Bennett, Merrill K. *Monograph in Economics: Population, Food, and Economic Progress.* Volume XXXIX, July 1952, Number 2. Houston, Texas: The Rice Institute.

Bickley, W. G., et al. *British Association Mathematical Tables, Volume X: Bessel Functions, Part II.* New York: Cambridge University Press, 1952. $11.00.

Blair, Morris Myers. *Elementary Statistics.* New York: Henry Holt and Co., 1952. $5.50.

Bogue, Donald J. *A Methodological Study of Migration and Labor Mobility.* Oxford, Ohio: Scripps Foundation, Miami University, 1952. Paper.

Bohmfalk, John F., Jr. *Chemical Investments: An Analysis of Research, Growth and Finances in the Chemical Industry.* New York: Chemonomics, Inc. (400 Madison Ave.), 1952. Paper. $5.00.

California, Bureau of Marine Fisheries. *The Commercial Fish Catch of California for the Year 1950 with a Description of Methods Used in Collecting and Compiling the Statistics.* Fish Bulletin 86. 1952. Paper.

Dun and Bradstreet, Inc. *Commercial Failures in an Era of Business Progress, 1900–1952.* Paper.

Duncan, Julian S., Staus, Harold O., and Nutt, Katharine. *Highway Finance in New Mexico.* Albuquerque: University of New Mexico, 1952. Paper. 50 cents.

Finney, D. J. *Statistical Method in Biological Assay.* New York: Hafner Publishing Co., 1952. $11.00.

Freund, John E. *Modern Elementary Statistics.* New York: Prentice-Hall, Inc., 1952.

Goldner, William. *Hours of Work.* Berkeley: University of California, Institute of Industrial Relations, 1952. Paper. 25 cents.

Gorter, Wytze, and Hildebrand, George H. *The Pacific Coast Maritime Shipping Industry, 1930–1948. Volume I, An Economic Profile.* Berkeley and Los Angeles: University of California Press, 1952. $2.50.

Goulden, C. H. *Methods of Statistical Analysis, Second Edition.* New York: John Wiley and Sons, 1952. $7.50.

Great Britain, General Register Office. *Cancer Registration in England and Wales. Third Year Recovery and Survival Rates.* Studies on Medical and Population Subjects, Supplement to No. 3. London: HMSO, 1952. Paper 1 s 3 d.

Inter-American Statistical Institute, Journal of. Estadistica. *Bibliography of Statistics Textbooks and other Teaching Material.* Washington, D. C., 1952. Paper.

Jambunathan, M. V. *Linear Estimation.* Bangalore: India Book Co., 1951.

McEntire, Davis, editor. *Inventory of Source Material and Research in Social Statistics, Pacific Coast, 1944–1950.* Berkeley: Social Science Research Council, Pacific Coast Committee on Social Statistics, July 1952. Paper.

McKinsey, J. C. C. *Introduction to the Theory of Games.* New York: McGraw-Hill Book Co., 1952. $6.50.

Machinery and Allied Products Institute. *Special Amortization of Defense Facilities.* Washington D. C.: MAPI (1200 18th St., N.W.), 1952. Paper. $1.00.

Maryland, University of, Bureau of Business and Economic Research. *Retail Prices and the Consumer Preference.* Volume Six, Number One. College Park, Maryland: 1952. Paper.

Milbank Memorial Fund. *Approaches to Problems of High Fertility in Agrarian Societies.* Papers presented at the 1951 Annual Conference. New York: 1952. Paper. $1.00.

National Bureau of Economic Research. *Conference on Research in Business Finance.* New York: 1952. $5.00.

Organisation for European Economic Cooperation. *The Internal Financial Situation.* Paris: 1952. Paper. 75 cents.

Phelps, Clyde William. *The Role of the Sales Finance Companies in the American Economy.* Baltimore, Maryland: Commercial Credit Co. (14 Light Street), 1952.

Spiegel, Henry William, editor. *The Development of Economic Thought: Great Economists in Perspective.* New York: John Wiley and Sons, 1952. $6.50.

Stein, Herman D. *Measuring Your Public Relations.* New York: National Publicity Council (257 Fourth Street), 1952. Paper. $1.25.

Suranyi-Unger, Theo. *Comparative Economic Systems.* New York: McGraw-Hill Book Co., 1952. $5.50.

Thompson, James H. *The Manufacturing Industries of West Virginia.* Morgantown: West Virginia University, Bureau of Business Research, 1952. Paper.

Tungate, L. A. *Financial Management for the Small Businessman.* Boston: Chapman and Grimes, 1952. $3.00.

United Nations. *International Tax Agreements, Volume III.* 1951 XVI 5. Paper. $4.00.

———. *Latin America, 1935–1949. A Selected Bibliography.* 1952. 1.12. Paper. $1.50.

———. *Statistics of National Income and Expenditure.* 1951 XVII 2. Paper. 50 cents.

United Nations Statistical Office. *National and Per Capita Incomes, Seventy Countries, 1949.* Statistical Papers Series E., No. 1. New York: 1952.

U. S. Department of Agriculture, Bureau of Human Nutrition and Home Economics. *Rural Family Living Charts.* Washington, D. C.: USGPO, October 1952. Paper.

U. S. Department of Commerce. *Better Population Forecasting for Areas and Communities.* Domestic Commerce Series No. 32. Washington, D. C.: USGPO, 1952. Paper. 25 cents.

Wilson, E. Bright, Jr. *An Introduction to Scientific Research.* New York: McGraw-Hill Book Co., 1952. $6.00.

World Health Organization. *Comparability of Statistics of Causes of Death According to the Fifth and Sixth Revisions of the International List.* Geneva. 1952. Paper. 50 cents.

Zimmerman, Carl C., and DuWors, Richard E. *Graphic Regional Sociology.* Cambridge, Mass.: The Phillips Book Store, 1952. Paper. $3.50.

RANDOM DIGITS (6001–6100)

From *A Million Random Digits*, to be published by The Rand Corporation, Santa Monica, California

Digits 1–6000 were published in Vol. 47, pp. 710–14 (December 1952).

77603	59467	58309	87834	57213
73258	73452	17619	56421	40725
03469	27635	56293	91700	04391
29718	86040	02596	01655	09918
45938	52403	94255	26351	46527
37510	14663	46581	61499	04844
23439	53633	94365	92560	12363
67317	43514	34261	08769	91830
45161	98588	37110	66538	39318
68224	71568	83950	45424	72453
94626	49465	46758	61630	97966
30246	02086	73750	01929	17165
23313	18256	57256	41659	39098
15626	44324	93119	32031	39608
19444	68219	40744	90043	93478
76537	46467	30942	14199	79418
12037	74558	16250	62308	85447
23982	29899	71594	13780	86219
75992	73445	01317	26743	71712
58044	06949	31176	96750	29776

1953 OFFICERS AMERICAN STATISTICAL ASSOCIATION

President	*President-Elect*	*Past President*
WILLIAM G. COCHRAN	HERBERT MARSHALL	ARYNESS JOY WICKENS

Vice Presidents

MORRIS H. HANSEN	WALTER E. HOADLEY, JR.	RENSIS LIKERT

Directors

W. J. DIXON	HOWARD L. JONES	JOHN W. TUKEY
MARGARET J. HAGOOD	WILLIAM R. LEONARD	RALPH J. WATKINS

Secretary-Treasurer

SAMUEL WEISS

Members of the Council

LUCILE DERRICK	HOWARD LEVENE	HARRY S. SCHWARTZ
A. ROSS ECKLER	WILLIAM G. MADOW	JOHN R. STOCKTON
MARTIN R. GAINSBRUGH	KENNETH E. MILLER	RUFUS S. TUCKER
EZRA GLASER	ALMARIN PHILLIPS	DAVID F. VOTAW, JR
WERNER HOCHWALD	ORVILLE B. RAILEY	F. M. WADLEY
A. J. JAFFE	J. A. RIGNEY	W. ALLEN WALLIS
MILDRED M. LAUDER		W. J. YOUDEN

A cumulative Index to Volumes 1–34, 1888–1939, may be obtained from the Office of the Secretary of the American Statistical Association.
Reprints of all articles published since 1945 may be purchased from the office of the secretary. Articles which have appeared during the past twelve months are usually available in quantity.